# TAKING THE NATURALISTIC TURN SERIOUSLY: A CRITIQUE OF NATURALIZED PHILOSOPHY OF SCIENCE

by

Philip Graham Ashton Good

Submitted in accordance with the requirements for the degree of PhD in Philosophy

Division of History and Philosophy of Science,
School of Philosophy

The University of Leeds

September 2003

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

## ACKNOWLEDGMENTS


I would like to thank the staff of the Division of History and Philosophy of Science, in particular Professor Steven French and Dr Graeme Gooday without whose support and advice I could not have completed this thesis.

In addition I would like to express my sincere gratitude to the Darwin Trust of Edinburgh for financial support and to Professor Peter Lipton of the University of Cambridge for his interest and feedback on chapter 4. I would also like to thank Dr Tarek Hayfa for helping me to appreciate the work of Richard Rorty and for valuable discussions of many of the theses presented here.

Finally, I would like to say thank you to my friends and family for their support and help. In particular I would like to acknowledge the support of Graham Good, June Good, Jeremy Good, Emma Alderdice, Allan Draper, Mary & Alfred Boulton, Phillip Darling and Gillian Magee.

University of Leeds

Abstract

# Taking the Naturalistic Turn Seriously: A Critique Of Naturalized Philosophy of Science

by Philip Graham Ashton Good

Division of History and Philosophy of Science, School of Philosophy

This thesis attempts to assess the viability of arguing for realism from a naturalistic perspective. It demonstrates that extant attempts to carry out this project fail to establish realism as a better explanation of science than various antirealist alternatives (particular attention is given to Laudan's naturalistic antirealism, van Fraassen's constructive empiricism and social constructivist accounts of science). In particular, it is shown that various attempts to refute Laudan's pessimistic induction are not successful. Further, it is argued that there is a good prima facie case for saying that naturalism and realism are not compatible philosophical positions. Accordingly, an attempt is made to develop a naturalistic account of science that is neither realist nor antirealist. Here it is argued that the main candidate for such a position, Arthur Fine's NOA, faces four major problems. In order to fix these problems I turn to Rorty's antirepresentationalist account of science and culture but conclude that as an attempt to develop and defend NOA, this account is inadequate in several important ways.

# Table of Contents



## Chapter 1

## What is Epistemological Naturalism?

# Chapter 2

# The Genesis Of Naturalistic Realism

# Chapter 3

## A Defence of Laudan's Pessimistic Induction: Part 1

# Chapter 4

## A Defence of Laudan's Pessimistic Induction: Part 2

# Chapter 5

# Naturalism, Realism and Scepticism

# Chapter 6

## Fine and the Four Problems of NOA

# Chapter 7

## Rorty's Antirepresentationalism: NOA's Ark or Neurath's Boat?

# Introduction

Like its sister discipline, naturalised epistemology, naturalised philosophy of science attempts to steer a course between traditional normative, a priori philosophy and more recent purely descriptive anti-normative projects. In doing so it attempts to honour Kuhn's (1970) seminal advice to philosophers that any account of science must take account of the actual practice and development of science. In this respect at least naturalised philosophers of science are in full agreement with sociologists of science who point out the importance of social and psychological factors in the development of a descriptively adequate account of science. However, in accordance with their melioristic vision, naturalised philosophers of science reject the claim that such descriptive adequacy is all we can hope for. Accordingly, they seek a way of marrying the sociologist's demand for empirical detail with the traditional philosopher's desire for normative import.

This general commitment to the development of a descriptive yet normative account of science is compatible with a wide variety of approaches, particularly in terms of just how descriptive one is prepared to be. Here there have typically been two sorts of attempt to naturalise the philosophy of science: a mild form ('weak scientistic naturalism') and a strong form ('strong scientistic naturalism'). Weak scientistic naturalism suggests that a properly descriptive philosophy of science should be based on the forms of reasoning and methodological rules that are characteristic of science itself. Thus, weak scientistic naturalists like Laudan (1977, 1984) attempt to show that philosophical claims about science can be treated as empirical hypotheses that can be tested against appropriate historical evidence. Strong scientistic naturalism endorses this point but, in addition, suggests that we should use particular scientific theories as substantive sources of empirical information. This latter commitment is particularly important to strong scientistic naturalists when the theories in question may have some bearing on the cognitive lives of scientists and/or the social organisation of scientific communities.

As far as carrying out the normative dimension of their project goes, naturalised philosophers of science have typically chosen to focus on the traditional issues of mainstream philosophy of science, e.g. whether or not science is progressive, the epistemic and ontological status of unobservable processes and entities, the rationality of science, etc. This thesis is concerned with weak and strong naturalistic attempts to establish realism as the best (scientific) explanation of science. In order to get a good understanding of exactly what it means to be a naturalist, chapter 1 discusses two attempts to explain naturalism in the philosophy of science due to Stump (1992) and Kitcher (1993). Here I argue that Stump's concept of a 'metamethod' fails to do justice to the varieties of naturalistic approaches in the philosophy of science. On a different note, I argue that although Kitcher can provide a perfectly adequate explanation of the *descriptive* features of naturalised philosophy of science he struggles when it comes to explaining the role of realism in such accounts. In chapter 2, I introduce Rosenberg's account of philosophical naturalism and support its claim that realism in naturalised philosophy of science is based on an the attempt to defeat three main antirealist rivals: Laudan's naturalistic antirealism, van Fraassen's constructive empiricism, and various 'constructivist' accounts provided by recent sociology of science.

Chapters 3 and 4 are one long argument against realist attempts to refute Laudan's pessimistic induction. Here I argue that recent attempts to do this by both naturalists (Psillos and Kitcher) and non-naturalists (Peter Lewis) fail to provide sufficient reasons to reject this long-standing argument for antirealism. In chapter 5, I turn to the other two major antirealist positions that preoccupy the naturalistic realist: van Fraassen's constructive empiricism and social constructivism. Here I combine a discussion of the available responses to these two antirealist accounts with a discussion of two fundamental problems that face any attempt to argue for realism from a naturalistic perspective: the problems of epistemic and methodological circularity. My suggestion is that when taken as a whole these arguments allow us to reach a rather surprising conclusion, namely that realism may not in fact be compatible with naturalism. I conclude that naturalism might be better off without realism, and vice versa.

Having established that there is a fundamental tension at the heart of naturalistic realism, chapter 6 introduces Fine's attempt to outline a position that is neither realist nor antirealist. Here I argue that the many critiques of Fine's account can be thought of in terms of four key problems. The suggestion is that a solution to these problems would provide us with a genuine alternative to realism and antirealism, a refuge for naturalists should they care to avoid this issue. In chapter 7 I conclude my analysis by assessing Rorty's antirepresentationalism as one possible way of achieving such solutions. However, I am forced to conclude that although Rorty can 'solve' the four problems of

NOA he does so at the expense of many of the minimalist virtues that made it seem attractive in the first place. I conclude with some suggestions concerning where we should look next in our quest for a naturalistic philosophy of science that is neither realist nor antirealist.

# Chapter 1

# What is Epistemological Naturalism?

## 1. Introduction

What does it mean to 'take the naturalistic turn'? In a recent book dedicated to this very question Callebaut tells us that, "the naturalistic perspective implies that *matters of fact* are as relevant to philosophical theory as they are relevant in science" (Callebaut 1993, p.1). This definition is a good starting point but it raises three obvious questions:

1. Which matters of fact do naturalists appeal to?

2. How relevant are matters of fact to philosophical theories?

3. Which philosophical theories can we naturalise?

In answer to the first question, most naturalists suggest that it is *scientific* matters of fact that we must appeal to when naturalising a particular philosophical theory. This presents the aspirant naturalist with the further question of which scientific theories or facts are most relevant for their purposes which in turn leads us on to the issue raised by the third question concerning which particular philosophical theories we wish to naturalise. Here we must note that there have been attempts to naturalize nearly all of the major sub-disciplines of philosophy, e.g. metaphysics, epistemology, philosophy of science, ethics, philosophy of mind, aesthetics, etc.

So, there are a huge variety of naturalistic positions depending on what matters of fact we choose to appeal to, how relevant we take these matters of fact to be, and which philosophical theory we want to apply them to. However, the aim of this chapter is primarily to investigate various attempts to naturalize epistemology and the philosophy of science. In section 2, I discuss Kitcher's attempt to explain epistemological naturalism suggesting that his account provides a better explanation of naturalized epistemology than that of Kornblith (1985). In section 3, I follow a suggestion of Stump (1992) and discuss naturalism in the philosophy of science in terms of a commitment to a particular 'metamethod.' Here I argue that although Stump's analysis can capture some aspects of naturalised philosophy of science it ultimately fails to provide a satisfactory account. In section 4, I attempt to provide a more nuanced account of some of the more important features of naturalised philosophy of science. I suggest that a good way of capturing the variety of approaches in this area is to think of them in terms of a distinction between weak scientistic naturalism and strong scientistic naturalism. In section 5, I return to Kitcher's account of epistemological naturalism and assess it as an account of naturalised philosophy of science, particularly in terms of the distinction introduced in section 4. I conclude that although Kitcher can account for its general structure we may have to turn elsewhere in order to fully appreciate the role that normativity plays in naturalised philosophy of science.

## 2. Epistemological naturalism

In his 1992 article, 'The Naturalists Return,' Philip Kitcher attempts to trace the development of contemporary epistemological naturalism. He begins by claiming that it should be seen as a return to an earlier approach to epistemology that was temporarily lost with the advent of modern analytical philosophy:

> Pre-Fregean modern philosophy was distinguished not only by its emphasis on problems of knowledge, but also on its willingness to draw on the ideas of the emerging sciences, to cull concepts from ventures in psychology and physics, later still to find inspiration in Darwin. (Kitcher 1992, p.54)

Kitcher is not alone in supporting this view; Callebaut (1993) has expressed a similar historical reading with respect to the philosophy of science:

> What has been going on in science studies in the last fifteen years can be best understood and made sense of by relating it to the *naturalistic turn* in the philosophy of science – actually the return of naturalism, after the interregnum during the most of this century of linguistically oriented philosophy. (Callebaut 1993, p.1)

So, for both Callebaut and Kitcher, the work of Frege is pivotal in rejecting a significant strain of naturalism in the history of philosophy. Not only did Frege replace epistemology with philosophy of language as the central concern of philosophy, he also rejected the use of psychology in this enterprise. Logic, not psychology or any other branch of science, became the proper tool of philosophical analysis.

## 2.1 Kitcher's traditional naturalism

According to Kitcher (1992) we can view the Fregean approach to philosophy as providing the foundation for analytical epistemology in terms of two crucial presuppositions:

1. A commitment to answering epistemological questions in a way that does not appeal to psychological facts about epistemic subjects.

2. The belief that philosophy is an a priori discipline based on the method of logical analysis where epistemological problems "reduce to questions of logic, conceptual analysis or grammar."[1]

(Kitcher 1992, p.56)

For Kitcher, the rise of epistemological naturalism in the latter half of this century can be understood as a rejection of these two presuppositions and a consequent return to our pre-Fregean ways.

Firstly, Kitcher explains how and why psychological facts came to be seen as necessary for philosophical discussions of epistemic justification. Within epistemology, arguments by Gettier (1963) showed that the analysis of knowledge in terms of justified true belief does not rule out certain cases where we could have such beliefs but could not be said to have knowledge. As Kitcher shows, this inevitably led to the introduction of psychology as an important resource for explaining the belief structure of epistemic subjects, thus breaking with the apsychologism of Frege:

> Analyses of the concept of knowledge (and, later, that of justification) were no longer confined to specifying the logical relations among propositions believed by the subject but could take into account the processes, including inevitably the psychological sub-processes, that causally generate states of belief. (Kitcher 1992, p.60)

Kitcher also notes the importance of developments in psychology that provided epistemologists with a new vocabulary for discussing the psychological lives of

[1] Of course, as Kitcher freely admits, these two presuppositions are intimately linked, for the injunction against psychology can be seen as a consequence of the more general commitment to a priori philosophy. I will have more to say about this issue later.

individual subjects. This, coupled with the nature of the responses to Gettier, explains the gradual rejection of the apsychologistic presupposition.

Kitcher identifies the second important factor in the development of epistemological naturalism as the rejection of the view that epistemological principles can be generated a priori. The main motivation for this, he argues, is Quine's attack on the notion of analyticity and Kuhn's work in the philosophy of science.[2] The former's influence cannot be disputed because in questioning the distinction between analytic and synthetic truths Quine thereby undermined the notion that anything could be known independently of experience, including epistemological principles. Kuhn's influence is less obvious because his work arose as part of a general move away from the approach of the logical positivist view of science. However, as Kitcher rightly claims, at least on one reading of his work we can view Kuhn as objecting to the a priori nature of the positivist's logic of science by "using the mismatch between the deliverances of methodologies for science and scientific practice to undermine our confidence in *a priori* pronouncements about how science ought to be done" (Kitcher 1992, p.70). I shall have more to say about Kuhn in the later sections but for now I want to concentrate on Kitcher's claim that the rejection of the two presuppositions leads to a single position, namely "traditional naturalism."

Traditional naturalism is Kitcher's attempt to define the core commitments of epistemological naturalism. He interprets this as being composed of four theses:

1. The central problem of epistemology is to understand the epistemic quality of human cognitive performance, and to specify strategies through whose use human beings can improve their cognitive states.
2. The epistemic status of a state is dependent on the processes that generate and sustain it.
3. The central epistemological project is to be carried out by describing processes that are reliable, in the sense that they would have a high frequency of generating epistemically virtuous states in human beings in our world.
4. Virtually nothing is knowable a priori, and, in particular, no epistemological principle is knowable a priori.

(Kitcher 1992, pp.74-76)

[2] See "Two Dogmas of Empiricism" in Quine (1951) and Kuhn (1970).

Theses (2) and (4) clearly follow from the critiques of apsychologism and the status of a priori epistemological principles respectively, but what about (1) and (3)? The answer to this question concerns the status of epistemology as a normative enterprise and it is this that explains Kitcher's own interest in naturalistic epistemology. Thesis (1) can be read as the view that although naturalized epistemology breaks with traditional epistemology in a number of important ways, it is still committed to the view that epistemology is a normative discipline. The job of naturalized epistemology is not only to tell us how in fact we arrive at our beliefs but also how we ought to arrive at our beliefs. Thesis (3) makes the further claim that the best way of doing this within a naturalistic framework is to focus on the reliability of beliefs. Thus "traditional naturalism" in Kitcher's sense includes a commitment to reliabilism (Goldman 1986).

In a moment I want to assess this account in terms of its ability to capture various naturalistic approaches in epistemology and the philosophy of science. However, before doing so I want to show that on its own Kitcher's definition of "traditional naturalism" is incomplete. As Kitcher is keen to point out, understanding contemporary naturalism requires an appreciation of the instability of its central claims:

> Traditional naturalists occupy an uncomfortable middle ground between earlier epistemologists and those who campaign for abandoning (or relativizing) normative projects. In the ensuing debates, each of the extreme positions can use its counterpart as a foil for denying the possibility of an intermediate position. Thus, post-Fregean epistemologists can attempt to oppose traditional naturalism by contending that it leads immediately to forms of relativism or scepticism that are unacceptable. Radical naturalists, by contrast, portray traditional naturalism as failing to break free from the errors of post-Fregean epistemology. (Kitcher 1992, p.77)

Kitcher goes on to show how this situation can be cashed out in terms of the modification of the four central theses of traditional naturalism. For example, the fact that epistemology is no longer regarded as a priori (thesis 4) can be read in a variety of ways. In particular, traditional epistemologists can prima facie accept this claim without thereby committing themselves to a fully naturalized epistemology based on empirical investigation. For saying that the tools of philosophical analysis do not have the a priori status once attributed to them does not automatically invalidate their use. According to this interpretation epistemology can continue as before with only a minimal reliance on the results of science.

However, Kitcher shows how problematic this weakened position is because thesis (3) tells us that in order to produce epistemological recommendations that are relevant to actual cognitive performance we must rely on empirical information about how reliable beliefs are formed. This creates a problem that faces the naturalist at every turn, the spectre of circularity. For if, as the naturalist suggests, we must rely on contingent empirical information to construct relevant epistemic recommendations where does the warrant for this information come from in the first place? Again Kitcher explains this situation very well:

> Because adequate epistemological principles emerge only late in inquiry – if at all – they must be based on a picture of nature obtained by using error-prone strategies. Consequently, the *apparent* information used in formulating our epistemic recommendations is likely to be misleading, with the result that *what we take to be* correct epistemic recommendations are infected with mistakes. (Kitcher 1992, p.79)

We shall see this complaint again in many different forms and there is a lot more to be said about it (see chapter 5) but at the present time its importance lies in the way it is used by the opponents of Kitcher's traditional naturalist.

If we accept that the circularity inherent in normative naturalism is bound to be vicious then it seems that the only kind of non-circular epistemic principles that will guide inquiry will be ones that are generated a priori. As Kitcher shows, for those who accept this argument there are two ways of evaluating its consequences for "traditional naturalism". Firstly, there is what one might call the conservative interpretation associated with traditional epistemologists who are keen to protect the normative status of epistemology. For if, as the argument suggests, nomativity requires an a priori source and naturalism questions the status of such sources, then so much the worse for naturalism. In terms of the four theses of "traditional naturalism" this would require the rejection of (3) and (4). Secondly there is what we might call the radical interpretation associated with philosophers who question the very possibility of normative epistemological projects. Radical naturalists accept the argument that only a priori epistemic recommendations can be genuinely normative but claim, with the traditional naturalist, that there can be no such recommendations. Therefore, they conclude that normative epistemology, and particularly the traditional naturalist's attempt to construct normative principles, is doomed from the outset.

I believe that the above sums up the core of Kitcher's attempt to locate and define contemporary naturalistic approaches. For the purposes of analysis I believe that the simplest and perhaps best way to look at Kitcher's account is diagrammatically, as a synchronic reconstruction of naturalism and its relationship to competing positions:

| | | | | |
|---|---|---|---|---|
| Post-Fregean Epistemology | ↔ | Traditional Naturalism | ↔ | Radical Naturalism |

The arrows try to capture the precarious position of traditional naturalism that results from its attempt to establish normative epistemic rules within a naturalistic framework. The value of this diagrammatic reconstruction of Kitcher's account will become clear in the later sections where it will be assessed in terms of its ability to capture naturalistic approaches in both epistemology and the philosophy of science.

### 2.2 The scope of traditional naturalism

Before attempting to assess Kitcher's account of naturalism I want to discuss a possible problem concerning the arguments of following sections. For although it may be legitimate to assess "traditional naturalism" as a description of naturalism in epistemology, it does not thereby follow that it must also serve as a description of epistemological naturalism in the philosophy of science. However, although Kitcher's "traditional naturalism" is primarily intended as a description of the status of naturalism in epistemology I believe he also intends it to serve as the epistemological foundation for naturalism in other branches of philosophy, particularly contemporary naturalized philosophy of science.

In order to see this it is necessary to attend to a possible ambiguity that arises from what one means by the term "epistemology." In a narrow sense "epistemology" is used to denote the particular academic discipline of contemporary epistemology, in a second much broader sense "epistemology" is used to denote any field that discusses epistemological issues. In the former sense what counts as epistemology is judged in terms of belonging to the particular academic discipline denoted by the proper name "epistemology." In contrast, the latter sense of "epistemology" denotes any work that is concerned with questions concerning knowledge *no matter which academic field they may occur in.* This may seem a rather trivial point but it does show why Kitcher's account of naturalism is more general than its concern with epistemology might indicate. For it is clear that Kitcher wants his account to explain naturalism in epistemology in both the narrow and broad senses of this term. This is obvious from the scope of Kitcher's article, which includes discussions of accounts in both sociology of science and the philosophy of science. These discussions only make sense if we view Kitcher as attempting to describe naturalism in epistemology in its broad sense – one that includes epistemology, sociology of science and philosophy of science.

This shows that Kitcher's account is to be viewed at a level of generality that transcends the concerns of contemporary epistemology in its narrow sense. However, even given this increased scope it is still far from clear whether or not we can push still further and ask if Kitcher is providing an account of naturalism that will apply to philosophical fields beyond those that are broadly epistemological. For there are areas such as philosophy of mind and ethics that do not necessarily take epistemological issues as their central concerns. It is less clear whether Kitcher's account is supposed to be applicable to these fields. Again we can represent this as follows:

Level 1: Naturalized epistemology (narrow sense) – naturalistic approaches in contemporary epistemology only.

Level 2: Naturalized epistemology (broad sense) - includes level 1 plus other subdisciplines concerned with epistemological issues, e.g. sociology of knowledge, philosophy of science, social constructivism, etc.

Level 3: Philosophical naturalism - includes levels 1 and 2, plus philosophical subdisciplines that are not primarily connected with epistemological issues, e.g. naturalized philosophy of mind, evolutionary ethics, etc.

I hope it is clear how these three levels relate to each other and why they help to capture the scope of Kitcher's account of naturalism. Level 1 is a subset of level 2, which in turn is a subset of level 3.

Kitcher's account not only applies to level 1, but also to level 2. Thus, in terms of the above diagram, we can see how Kitcher applies his framework to understand epistemological issues in science studies:

| | | | | | |
|---|---|---|---|---|---|
| General Framework: | Post-Fregean epistemology | ↔ | Traditional naturalism | ↔ | Radical naturalism |
| Level 1: | Traditional epistemology | ↔ | Naturalized epistemology | ↔ | Anti-normative epistemology |
| Level 2: | Logical Positivism | ↔ | Naturalized philosophy of Science | ↔ | SSK, social constructivism |

This shows that the appropriate questions to ask of Kitcher's account are not about whether or not he gives a good account of philosophical naturalism (level 3), for this is not what he is after; rather it is about the adequacy of his account for explaining epistemology in both the narrow and broad senses of this term (level 1 and 2). This raises two important questions concerning Kitcher's treatment of naturalism in this field. Firstly, is "traditional naturalism" a good description of naturalism in epistemology? Secondly, is "traditional naturalism" a good description of naturalism in the philosophy of science? Given my concern with understanding naturalism in the latter of these two areas I will devote the bulk of this chapter to answering the second of these questions. However, before doing so I will briefly attempt to answer the first question concerning the adequacy of Kitcher's account for understanding naturalism in epistemology.

## 2.3 Traditional naturalism in epistemology

As far as the first question is concerned I think we must say that Kitcher gives a very good account of the current state of naturalized epistemology. Indeed Kitcher's definition of "traditional naturalism," and its relations to competing epistemic approaches, shares many features with Kornblith's introduction to *Naturalizing Epistemology* (1985). In fact, I propose that a good way of assessing Kitcher's contribution here is to compare it with Kornblith's excellent classification of naturalistic epistemologies.

Kornblith attempts to show how naturalism in epistemology results from how one conceives the relationship between the following three questions:

1. How ought we to arrive at our beliefs?
2. How do we arrive at our beliefs?
3. Are the processes by which we do arrive at our beliefs the ones by which we ought to arrive at our beliefs?

The "traditional view" according to Kornblith is to assign question (1) to philosophers and question (2) to psychologists. Then the answer to question (3) will come from a comparison of the answers provided by these two separate fields. In contrast, Kornblith believes that the best way of understanding naturalized epistemology is to see it as based on the claim that question (1) cannot be answered independently of question (2). He speaks of this in terms of a *replacement* of epistemological questions by psychological questions.

Of course, as Kornblith realizes, how far one pushes the replacement thesis is the key to understanding the variety of naturalized epistemologies currently available. Broadly speaking this issue results in two possible approaches, the strong replacement thesis where:

> Psychological questions hold all the content there is in epistemological questions. On this view psychology replaces epistemology in much the same way that chemistry has replaced alchemy. (Kornblith 1985, p.6)

And the weak replacement thesis:

> In this view psychology and epistemology provide two different avenues for arriving at the same place. Psychology may replace epistemology because the processes psychologists identify as the ones by which we do arrive at our beliefs will inevitably turn out to be the very processes epistemologists would identify as the ones by which we ought to arrive at our beliefs. (Kornblith 1985, p.6)

Prima facie, we now have a scheme that is comparable to Kitcher's, for the weak and strong replacement theses in Kornblith's account look very much like traditional and radical naturalism in Kitcher's. This suggests the following schema:

| | | | | |
|---|---|---|---|---|
| Traditional View | ↔ | Weak Replacement Thesis | ↔ | Strong Replacement Thesis |
| ↑↓ | | ↑↓ (?) | | ↓↑ (?) |
| Post-Fregean Epistemology | ↔ | Traditional Naturalism | ↔ | Radical Naturalism |

Clearly the identification of Kornblith's "traditional view" with Kitcher's "Post-Fregean epistemology" will be unproblematic. The important questions concern the identification of the weak and strong replacement thesis with traditional and radical naturalism respectively. Are they the same or does Kitcher introduce a new (and possibly improved) way of looking at the status of naturalized epistemology?

Although Kitcher and Kornblith's accounts look similar there is a significant sense in which they differ. For although they both acknowledge the existence of a conservative and a more radical position within naturalism they do so for different reasons. Kornblith's focus is on the status of epistemology as an autonomous subject, in other words he is interested in the disciplinary effects of epistemological naturalism. This is why he talks in terms of replacement. Kitcher, in contrast, takes as his primary focus normativity and how the rejection of the a priori status of epistemic principles affects this issue. This is not to say that the autonomy of epistemology is of no interest to Kitcher, rather that he does not believe that it is the crux to an understanding of epistemological naturalism. This difference becomes clear when Kornblith discusses the a priori status of epistemology:

> Most naturalistic epistemologists will reject the suggestion that anything is knowable a priori. I do not wish to enter that debate here. Rather I will argue that the issue of a priority is a red herring. Whether the answers to epistemological questions can be known a priori in principle, epistemologists would do well to consult psychologists in practice. (Kornblith 1985, p.10)

At this stage we can recall from above the importance of apriority for Kitcher. The rejection of the possibility of a priori knowable epistemic principles was the very point that explained the precarious position of "traditional naturalism" between old school epistemology and radical anti-normative naturalism. Thus, apriority cannot be a red herring for Kitcher; it is crucial to understanding contemporary epistemic naturalism.

Having established that Kitcher's account of epistemological naturalism is significantly different from Kornblith's the question arises as to which provides the better description of naturalism in epistemology. For a number of different reasons I believe that it is Kitcher who we should favour in this respect. Firstly, Kitcher develops a sophisticated definition of what has become the "received view" of naturalized epistemology in terms of the four theses of "traditional naturalism." Kornblith does identify some of the issues these theses concern but he lacks a developed framework for discussing them. Not only can Kitcher's account tell us exactly what most epistemologists would call naturalized epistemology, he can also show in detail how other positions result from this core definition. Secondly, and perhaps most importantly, by focussing on the problem of circularity rather than questions of autonomy, Kitcher is able to draw attention to a very important feature that is absent from Kornblith's account. Namely that "traditional naturalism" attempts to steer a course between two competing extremes. As Rosenberg notes, this feature is not merely confined to epistemology:

> Naturalists in most of the subdisciplines of philosophy have viewed their project as one that seeks a reasonable compromise between competing extremes while doing as much justice to the extreme views as they deserve. (Rosenberg 1996, p.2)

Kornblith's account of the relationship between the weak replacement thesis and the competing extremes of the traditional view and the strong replacement thesis fails to capture this feature.

I hope it is now clear why Kitcher's account of epistemological naturalism is preferable to Kornblith's. The purpose of this detour was to assess the adequacy of "traditional naturalism" as a description of naturalism within epistemology. We have seen that not only does Kitcher offer an account that captures most of the important features of naturalism in epistemology; he also improves upon earlier work in this field. I now return to the second question that concerns the adequacy of Kitcher's account of epistemological naturalism: Is "traditional naturalism" a good description of naturalism in the philosophy of science?

Here I want to suggest that establishing the adequacy of Kitcher's account hangs on its ability to capture two important features of naturalized philosophy of science. The first of these concerns the applicability of Kitcher's account to the naturalistic methodology that underpins many contemporary accounts of science. The second concerns its ability to capture the variety of normative claims that characterize these accounts. The task of the rest of this chapter will be to provide the evidence necessary to address the first of these concerns.

## 3. Naturalistic 'metamethods' in the philosophy of science

In this section I focus on a variety of naturalistic philosophers who self-consciously attempt to situate their projects somewhere between the extremes of the logical positivists and their historicist successors. Without exception such philosophers are keen to hang on to what they see as the common core of these disparate positions, namely an interest in providing a general (or global) account of the scientific enterprise that can explain and perhaps justify its prevalent features. This "global" perspective coupled with a renewed interest in naturalism leads such philosophers to what they hope are new and interesting answers to age old questions concerning science, i.e. Is science progressive? What is the status of the unobservable entities postulated by scientific theories? What is the appropriate methodology for science? What is the nature of scientific rationality?

For reasons of space I have chosen to focus on four contemporary attempts to provide naturalistic answers to these questions: Kitcher (1993), Giere (1985, 1988), Hull (1988) and Laudan (1977, 1984). My reason for the selection of these accounts rather than others is that I believe they constitute the strongest representatives of certain trends within the development of naturalized philosophy of science in the last twenty years. Following a suggestion of Stump (1992), I will discuss these accounts in terms of the claim that they are best understood as being based on a particular 'metamethod.'

### 3.1 A plurality of metamethods

Although naturalism implies that empirical information must be relevant to the study of epistemological practice, including science, this does not in itself suggest which science or sciences should form the basis of a suitably naturalized account. To be sure the selection of relevant empirical information does restrict us to sciences that concern themselves with the functioning of human beings, but this won't get us very far.[3]

[3] Some philosophers of science have even attempted to naturalize their subject by using results from scientific fields that are not directly concerned with the functioning of human beings. For example,

Sciences that take human beings as their subject matter will include fields as diverse as psychology, sociology, evolutionary theory, economics, social anthropology, etc. The question for the naturalist is not only about which of these disciplines to draw on but also which particular theory within each discipline to use. Stump (1992) suggests that most naturalists in epistemology address this problem in a very reductionist manner:

> Epistemological naturalists have tried to find a discipline to be the methodological model for epistemology, generally some version of psychology – behaviourist for Quine (1969), and cognitive for Alvin Goldman (1986). (Stump 1992, p.457)

For the naturalized epistemologist the privileging of psychology is not particularly problematic because their concern with the justification of the knowledge of individual subjects immediately suggests the relevance of psychological theories of cognitive functioning. However, as the above quote suggests, there still remains the problem for the naturalized epistemologist of which particular theory of cognitive functioning is the right one to use.

For the naturalist concerned with explaining science this problem just gets worse because, unlike naturalized epistemology, there seems to be no criterion of selection beyond the suggestion that one use sciences that are relevant to the study of human beings. As suggested above, this leaves the naturalized philosopher of science with a bewildering array of sciences on which she can draw when constructing her account. Given this problem is particularly acute for the naturalized philosopher of science one might expect to find naturalized accounts of science based on a wide variety of scientific disciplines. However, Stump (1992) argues that, following the lead of naturalized epistemologists, naturalized philosophers of science have typically chosen to base their accounts on one particular scientific discipline or, as Stump puts it, one particular 'metamethod':

> Naturalists in the philosophy of science have followed suit and attempted to reduce methodology to one metamethod. This metamethod might again be psychology, as for Ronald Giere (1988, 1989), history of science, as for Larry Laudan (1984, 1987), evolutionary biology, as for Donald Campbell (1974), or sociology, as in the work of Steve Fuller (1988). (Stump 1992, p.457)

---

Thagard (1988, 1992) has attempted to formulate a naturalized account of science based on theories in computational science.

For present purposes, questions concerning the validity of this move will be sidelined. Instead, I want to look more closely at some of the naturalized accounts mentioned here (and some that are not) to see how they employ results from these fields. My treatment will be thematic, introducing various naturalized accounts of science based on the particular metamethod they privilege. I begin with psychology.

### 3.1.1 The role of psychology

One of the earliest programmatic statements for the development of naturalized philosophies of science due to Giere suggested that psychology and related disciplines would have to play a crucial role:

> Neither foundationism nor metamethodology can break the circle or provide the norms needed to defeat relativism. This hardly proves that there is no way to achieve these ends. It does, however, provide some motivation for seeking to understand how a naturalized philosophy of science might fruitfully be pursued. I would suggest that evolutionary theory, together with recent work in cognitive science and the neurosciences, provides a basis for such an understanding. (Giere, 1985, p.339)

Following up on this suggestion, Giere (1988) argues for a naturalized account of scientific practice and rationality based around the notion of scientists as limited cognitive agents interacting with their environment.[4] In doing so he attempts to utilise many findings from cognitive science that concern the actual functioning of human agents. As the above quote indicates, this approach is also strongly evolutionary in that such cognitive abilities are acknowledged to be the contingent products of natural selection. I will leave this aspect of Giere's approach for a later section, for the moment all we need note is that, as Stump suggests, Giere privileges psychology as the most relevant source for constructing a naturalized account of science.

An emphasis on the importance of cognitive psychology in naturalizing the philosophy of science is also evident in Kitcher (1993). Here Kitcher argues that we should view developments in scientific practice and theory in terms of the following schema:

| | | | | | | |
|---|---|---|---|---|---|---|
| Social and Asocial Interactions | ↔ | Changes in Cognitive States and Processes | ↔ | Changes in Individual Practices | ↔ | Changes in Consensus Practices |

This framework shows that Kitcher views the cognitive life of individual scientists as the key to understanding the development of science. Changes at the abstracted levels of both individual and consensus practice are ultimately dependent on changes at the level of individual cognitive states and processes. Like Giere, Kitcher makes it quite clear that only a naturalistic understanding of this level will do:

> Science is not done by logically omniscient lone knowers but by biological systems with certain kinds of capacities and limitations. At the most fine-grained level, scientific change involves modifications of the cognitive states of limited biological systems. (Kitcher, 1993, p.59)

Again, as with Giere, we can see the importance here of using our evolutionary development as justification for invoking psychological theories of (evolved) cognitive functioning.

However, this kind of argument won't settle which of the many scientific theories of cognitive functioning we should employ. Kitcher is fully aware of this problem but privileges theories that concentrate on the representational nature of cognitive performance. Even though this still leaves a number of possible alternatives Kitcher is prepared to go with one particular approach:

> There is no one psychological consensus about the nature of representations. Following *one* major approach in cognitive psychology and artificial intelligence, I shall assume that mental activity often involves the storage and retrieval of propositions, and that we think we can think of such processes on the analogy of the inscription, reading, and rewriting of statements in boxes. (Kitcher, 1993, p.63)

In order to give this proposal some structure Kitcher bases his analysis on a model of cognitive functioning drawn largely from the work of Anderson (1983). Broadly speaking this model distinguishes between items that are active in working memory and those that are stored in declarative memory, long-term goals and procedural memory[5].

---

[4] An earlier attempt to develop this kind of account can be found in De Mey (1982).

[5] For a detailed picture of this model see Kitcher 1993, p.64.

Kitcher acknowledges that acceptance of this theory is tentative at best, very much along the lines of a scientific hypothesis.

Further details of Kitcher's naturalized account will de discussed in the next chapter, for the moment all we need note is that both Kitcher (1993) and Giere (1985, 1988) use evolutionary theory as a basis for privileging cognitive psychology as perhaps the single most important scientific theory for developing a naturalistic understanding of science.

### 3.1.2 The role of evolutionary theory: multiple understandings

So, both Kitcher (1993) and Giere (1988) appeal to evolutionary theory in order to underpin their cognitive accounts of science. As Bradie (1994) has noted, this constitutes a familiar argument for naturalism:

> - Human beings, as the products of evolutionary development, are natural beings. Their capacities for knowledge and belief are honed by evolutionary considerations. As such, there is reason to suspect that knowing, as a natural activity, can and should be treated and analysed by the methods of science. On this view, there is no sharp division of labor between science and epistemology. In particular, the results of particular sciences such as evolutionary theory and cognitive psychology are deemed relevant to the solution of epistemological problems. (Bradie 1994, p.453)

Most epistemological naturalists either implicitly or explicitly accept this argument,[6] so we might be tempted to conclude that it captures the significance of evolutionary theory for naturalized philosophy of science. However, there is a major tradition within epistemological naturalism that goes well beyond this basic commitment to evolutionary theory. This tradition is commonly referred to as evolutionary epistemology.

Like "naturalism," the term "evolutionary epistemology" covers a multitude of sins. However, all the approaches it is deemed to apply to share a much stronger commitment to the use of evolutionary theory than that found in the accounts of philosophers like Kitcher and Giere. All Kitcher and Giere say is that if cognitive abilities are acknowledged to be the (highly contingent) products of natural selection then we cannot hope to understand them without taking account of their status as products of this process. Here the appeal to evolutionary development is merely a prelude to the development of a cognitive account of science. Following Bradie (1994) we might say that the cognitive accounts of Kitcher and Giere are at most indirectly motivated by evolutionary considerations. As we might expect the appeal to evolutionary theory in the accounts of evolutionary epistemologists is far more direct. However, we must be careful here because at least on one understanding of the term even Kitcher and Giere will count as evolutionary epistemologists. For as Bradie has shown there are two distinct programmes that go by this name:

> One is the attempt to account for the characteristics of cognitive mechanisms in animals and humans by a straightforward extension of the biological theory of evolution...The other program attempts to account for the evolution of ideas, scientific theories, and culture in general by using models and metaphors drawn from evolutionary biology. (Bradie 1994, p.454)

Bradie labels the first of these programmes EEM and the second EET. It is clear now why Kitcher and Giere will in one sense count as evolutionary epistemologists, for although they are not directly concerned with developing the EEM programme they do accept its main claims concerning the origin of cognitive structures. However, this is where they stop, for it is clear that Kitcher and Giere have no time for the second kind of project characteristic of evolutionary epistemology – the EET programme.

The EET programme in evolutionary epistemology is far more controversial than EEM. Almost anyone who accepts the results of evolutionary theory can acknowledge that some form of the EEM programme is bound to be correct. The same cannot be said of the attempt to apply evolutionary models and metaphors to the development of knowledge. It is one thing to say that our cognitive structures are the product of the process of natural selection and quite another to claim that the development of human knowledge might similarly be analysed in terms of this mechanism. With respect to analysing the growth of science Bradie shows that there are two important ways of developing this suggestion:

> Some of these attempts involve analysing the growth of human knowledge in terms of evolutionary (selectionist) models and metaphors (e.g., Popper 1968, 1972; Toulmin 1972; Hull 1988). Others (e.g., Ruse 1986; Rescher 1977) argue for a biological grounding of epistemological norms and methodologies but eschew

[6] In fact we can find this kind of conclusion not only in Kitcher's work on the philosophy of science but also in his earlier attempt to understand epistemological naturalism (Kitcher 1992, p.58).

> selectionist models of the growth of human knowledge as such. (Bradie 1994, p.454)

The proponents of the first of these projects agree with Kitcher and Giere that the direct application of evolutionary theory is limited to what it can tell us about the evolution of cognitive structures. However, unlike Kitcher and Giere, philosophers of science like Hull (1988) suggest that evolutionary theory may be relevant in a further *analogical* sense. Roughly speaking such approaches suggest that it is useful to think of the development of scientific knowledge in terms of the mechanism of natural selection. This project attempts to find analogies for evolutionary concepts like 'gene' and 'variation' in science, using them to show how science can be viewed as a process of the selection of successful scientific theories.[7]

Proponents of the second kind of approach reject this kind of analogical move. Instead they try to show that the methodologies and norms that guide scientific inquiry are the product of our evolutionary history. They thus argue for the direct application of evolutionary theory in a much stronger sense than the analogical theorists. For not only do they argue that our cognitive structures are the products of evolution (the minimal claim accepted by Kitcher, Giere and the analogical theorists), they also claim that scientific methodology itself is the product of natural selection. This latter claim is different from, but surely as controversial as, the claims of the analogical theorists.[8]

I hope the above illustrates the multiple roles evolutionary theory plays within a variety of naturalistic approaches; particularly those that attempt to explain the development and practice of science. The validity of some of these 'evolutionary' approaches seems highly suspect to say the least, yet it is not my aim here to show why this might be the case. In terms of Stump's treatment we might be tempted to conclude that both of the projects found within Bradie's EET programme privilege evolutionary theory as their preferred 'metamethod'. However, I want to suggest that this conclusion is a little too simplistic. Whereas as it might be fair to say that the appeal to evolutionary theory to explain the origins of scientific methodology constitutes a 'metamethod' (in the sense that this theory provides the basis for naturalistic theorizing), the same conclusion does not appear to capture the very different kind of appeal to this theory found in analogical approaches. In analysing the growth of

[7] As Hull (1988) notes, one should be careful not to misinterpret this view as the absurd suggestion that the selection against unsuccessful scientific theories is a result of the death of the scientists who support them.

knowledge these kinds of approaches do not appeal to evolutionary theory in a strictly scientific sense. Rather analogical theorists use it as a resource for a particular kind of theory structure (i.e. natural selection) that they believe can be usefully applied to the study of scientific development. Referring to both of these appeals to evolutionary theory as a 'metamethod' seems to be misleading.

Similar difficulties arise with respect to the cognitive accounts of Kitcher and Giere, for interpreting these as based on the metamethod of psychology tells us nothing about their commitment to evolutionary theory discussed above. This suggests that there may be something wrong, or at least missing, in thinking of naturalistic approaches in terms of metamethods. There appears to be a lot more subtlety to many naturalistic approaches than this concept can capture. As we shall see in the next section, this complaint is particularly relevant with respect to at least one of the analogical approaches discussed above, namely Hull (1988). Here we find an approach that marries the analogical use of evolutionary theory with a strongly sociological account of the functioning of science. However, in the absence of an alternative account of what is going on here I will continue the discussion on the understanding that it is sociology here that functions as a 'metamethod' in accounts like Hull's.

### 3.1.3 The role of sociology

In the last section we saw that Hull's commitment to an evolutionary approach can be viewed along the lines of developing the analogy between the evolution of biological species and the growth of knowledge. In fact, Hull argues that his account is evolutionary in a much stronger sense than this analogical characterisation allows for. He claims that the growth of scientific knowledge is not merely analogical to the evolution of scientific knowledge but is in fact an instantiation of the same kind of general selection process. This has the effect of reducing both of these processes to the same epistemological status. The only difference being that we currently have a much better understanding of the operation of the selection mechanism in biology. Hull's aim is to redress this imbalance by showing how the selection mechanism is equally applicable and fruitful when applied to the growth of knowledge.

In order to cash out this claim Hull attempts to show that science can be viewed along the lines of a selection process. Hull makes it quite clear that, in contrast to

---

[8] Here I shall not be concerned with the merits of these accounts. For further details see Bradie (1994).

Kitcher and Giere, his account of this process will be based on its social structure rather than psychological theories of epistemic subjects:

> Even in the context of science, I pay little attention to the psychological makeup of scientists, concentrating instead on their social organization. For the questions I am addressing, the psychology of the individual actors is less important than the structure of the communities in which they function. (Hull 1988, p.27)

Based on this claim Hull attempts to construct a sociological account of science that will explain some of its most important features:

> The System of cooperation and competition, secrecy and openness, rewards and punishments that have characterized science from its inception is both social and internal to science itself. The conceptual development of science would not have the characteristics it has without this social system. (Hull 1988, p.3)

Hull has much to say about the importance of various norms in determining the social structure of science. However, for present purposes the important point to note here is that Hull's attempt to outline such an account is crucially dependent on the sociological theories of Merton (1973).[9]

### 3.1.4 History of science as a metamethod

Since the publication of his 1977 work, *Progress and its Problems*, Larry Laudan has been at the forefront of the attempt to develop and defend a naturalized perspective within the philosophy of science. Although Laudan has since rejected many of his earlier claims about rationality and progress in science (particularly his insistence that problem solving is a transcendent goal of scientific practice – see Laudan 1984) he has retained the belief that a naturalistic philosophy of science remains the best, and perhaps only way to establish these concepts as genuine features of science. As we shall see this is particularly true of Laudan's continued insistence that normative accounts of science must be judged in terms of their ability to capture important features of past episodes in the history of science.

In *Progress and Its Problems* Laudan made his now famous (or infamous) claim that in order to achieve an accurate account of progress and rationality, scientific theories must be evaluated in terms of their "problem-solving effectiveness." Laudan

[9] For an alternative attempt to construct a normative sociology of science see Fuller (1988).

makes it quite clear early on that his account is to be judged in terms of naturalistic criteria:

> If science is a rationally well-founded system of inquiry, then it is only right and proper that we should emulate its methods, accept its conclusions, and adopt its presuppositions. (Laudan 1977, p.2)

Here we can see Laudan appealing to the success and power of science to underpin the methodology behind his philosophical account of science. For given that Laudan's whole motivation for writing *Progress and Its Problems* was to show that science *is* a rationally well-founded system of inquiry, the implication is that we should emulate its methods and procedures in philosophy of science.

But what would such an emulation of scientific method in the philosophy of science consist in? For Laudan (1977) the answer lies in rethinking the relationship between the normative practice of philosophising about science and its empirical counterpart – the history of science. As strange as it may seem to us now, many of Laudan's predecessors viewed this relationship as almost nonexistent because of the widespread acceptance in philosophical circles of the divide between matters of fact and value. In the philosophy of science this divide, policed by proponents of the naturalistic fallacy argument of G.E. Moore (1903), led to the following kind of view:

> History is irrelevant to the philosopher because he is not concerned about what science has been, but rather how it should be. Philosophy is irrelevant to the historian because it is not his job to make normative judgments about the figures he studies. (Laudan 1977, p.156)

Of course, anyone familiar with developments in philosophy of science in the last 40 years knows that this view is now largely out of favour, Kuhn (1970) in particular being instrumental in bringing about its decline. However, for Laudan (1977), Kuhn's insistence on the relevance of history of science for the construction of normative accounts of science merely provided the first steps towards a fully-fledged naturalism. Thus, for Laudan, the history of science is not just relevant to normative claims about science; it is the data against which such claims must be tested. It is important to note here just how strong this claim is, for according to this view not only do philosophical theories about science lose their previously privileged a priori status, they are reduced to the same empirical level as scientific theories themselves.

This has important consequences for Laudan's views on the philosophy of science. For not only does it require him to test his own problem solving account against various episodes in the history of science, it also rules out a number of competing approaches within the philosophy of science. The former will be discussed at length in the next section, as for the latter the most important approach ruled out by Laudan's commitment to testing theories of science is the rational reconstructionism advocated by Lakatos (1978). For Laudan this approach is inadequate because its reconstruction of past episodes in the history of science in order to make them more rational is no better than ignoring the history of science altogether. He sums up his objection in the following way:

> If the philosopher would learn something from history, he must make himself a servant to it – at least to the extent of dealing with actual cases and actual beliefs. (Laudan 1977, p.170)

In the rest of this section I want to show that this naturalistic aspect of Laudan's work has remained constant throughout many significant shifts in his views about science, in other words he has continued to be the faithful servant of history he set out to be.

In his 1984 work, *Science and Values*, Laudan continues the discussion of the role of cognitive values in science that had he had begun in *Progress and its Problems*. As in his earlier work, Laudan remains convinced that neither "old wave" logical empiricism, nor the "new wave" philosophy of science of Kuhn and Feyerabend, provides us with an adequate theory of scientific rationality. However, in his later work, Laudan (1984) no longer seems to think that the identification of a transcendent goal for science (i.e. problem solving) will be sufficient to move us beyond the errors of these conflicting approaches. Instead he poses a new kind of problem for the development of a unified theory of scientific rationality. For Laudan (1984) this problem is to provide a theory of rationality that will be able to capture the virtues of both of these approaches.

Laudan begins his attempt to provide such a theory by noting that traditional approaches to consensus-formation rely on a particular model of rationality:

> The best-known contemporary solution to the problem of consensus formation in science involves postulating what I call the hierarchical model of justification, although it is perhaps more commonly known as the theory of instrumental rationality. Proponents of this model typically envisage three interrelated levels at which, and by means of which, scientific consensus is forged. (Laudan 1984, p.23)

The three levels alluded to here are intended to provide an exhaustive classification of the kinds of debate we might expect to find in science ranging from debates over disputed facts to higher-order disagreements over the proper goals of science. This gives us the following hierarchical structure:

1. Factual Level – including claims about theoretical or unobservable entities.

2. Methodological Level – the rules of evidence and empirical support.

3. Axiological Level – the aims or goals of science.

Laudan believes that the only way to arrive at an adequate account of scientific rationality is to modify our views on the supposed interconnection between these levels of justification. The details of this account we will be discussed at length in the next chapter. Here I merely want to show how Laudan deploys certain historical examples in an attempt to test his claims about consensus-formation in science.

This feature of Laudan's approach is most obvious in his discussion of disputes over the manifest goals of science. Here Laudan wants to argue that, contrary to the views of proponents of the hierarchical model, there is the possibility of resolving disputes that occur at the axiological level of inquiry, i.e. disagreements over the proper aims and goals of science. In order to establish this Laudan provides us with an illustrative example from the history of science. Here he begins by noting that for many 19$^{th}$ century scientists it was important that we should avoid theories that postulated unobservable entities. However, due to various shifts in the scientific theories of the time this goal became more and more difficult to reconcile with the assumptions these theories made about the existence of unobservable entities. Laudan is able to show that the perceived success of these theories led, in time, to the gradual rejection of this methodological restriction on the appeal to unobservable entities.

For our purposes the details of this example are not particularly important. Instead the crucial point to note about Laudan's appeal to historical evidence here is that, once again, it emphasises Laudan's commitment to naturalism. In fact Laudan's example is particularly interesting because it makes the case for naturalism in two different ways. Firstly, Laudan's example shows that, as a matter of historical fact, scientists have been able to modify their goals by judging them against the character of scientific theories. This not only rules out the view that the pursuit of goals is purely a 'matter of taste or fashion,' but also shows that the resolution of axiological disputes is, at least sometimes, a matter for empirical investigation. In this way Laudan is able to argue that scientist's judgments concerning the pursuit of particular normative values are intimately connected to descriptive considerations at the level of theories. Secondly, and perhaps more importantly, the use of an example to establish the possibility of consensus-formation at the axiological level again reflects Laudan's naturalistic belief that meta-level theories of science must be judged in terms of their ability to capture important features scientific practice[10]. In other words, Laudan believes that his own normative account of science must be tested in terms of its descriptive adequacy.[11]

### 3.2 Is history of science a metamethodology?

When viewed from the perspective of the other accounts discussed in this section, there is something rather peculiar about Laudan's naturalism. For, unlike Kitcher, Giere and Hull, what is absent from Laudan's discussion is any kind of appeal to particular scientific theories to ground his account of science. Instead we find a general commitment to emulating the methods of science, particularly the testing of theories with empirical evidence. Stump's claim is that this is because, instead of opting for a scientific metamethodology (e.g. cognitive science, evolutionary theory, etc.), Laudan instead opts for the history of science. But what can it mean to say that history of science is functioning as a metamethodology here? Is history of science the kind of thing that can function as a metamethodology?

My worry here is that in trying to offer a generalization about naturalized philosophy of science Stump has confused Laudan's use of the history of science with the kind of scientistic naturalism that characterizes the approaches of Kitcher, Giere and Hull. As we have seen the latter kinds of accounts are united in the sense that they all appeal to science as a resource for constructing descriptively accurate accounts of science. Although I have given reasons to suggest that the concept of a metamethodology might not capture the complex nature of this appeal (e.g. Hull's appeal to both evolutionary theory and sociology) I still believe that it can help us to understand what is going on when scientific theories are employed in this context. For

[10] Doppelt (1986, 1990) and Siegel (1990) have argued that Laudan's arguments for consensus-formation at the axiological level lack a naturalistic foundation. However, Freedman (1999) has shown that these objections are based on a failure to appreciate two senses of naturalism that run through Laudan's work.

example, in calling Giere's use of cognitive science a metamethod Stump is quite rightly alluding to the importance of this scientific field for Giere's analysis of scientific representation and decision-making procedures.

However, I do not believe that Stump's concept of a metamethodology has the same kind of heuristic value with respect to Laudan's use of the history of science. In fact, I want to claim that it is positively misleading because if we allow that history of science is a metamethodology for Laudan then we will have to say exactly the same thing about the other accounts discussed in this section. For as I will show in the next section, all of the accounts considered above rely on empirical evidence as a resource for constructing and testing their claims about science. But, of course, this undercuts Stump's claim that we can distinguish Laudan's account on the grounds that it uses history rather than a particular science as its metamethodology. Equally, in the case of accounts that rely on the results of particular sciences, we will be left with the conclusion that there are two meta-methodologies at work here; one of them reflecting the appeal to a particular science, the other reflecting the use of historical evidence.

I believe that this should lead us to conclude that, pace Stump, the use of historical evidence in naturalized accounts of science is not the same as the attempt to ground such accounts on the results of a particular science. Clearly then we need an alternative account that can explain what the relevant differences between these two features are. The development of such an account will be the aim of the next section.

## 4. Taking the descriptive turn seriously: The role of empirical evidence.

The purpose of this section is to show that the methodology that underpins Laudan's naturalism can also be found in the accounts of Kitcher (1993), Giere (1988) and Hull (1988). In addition to using the results of particular scientific theories these accounts follow Laudan in placing great importance on the use of methodological procedures traditionally associated with science, e.g. hypothesis formation and testing. However, although these accounts share a commitment to these procedures they interpret this feature of naturalism in a variety of different ways. As I will now show this is most obvious in their use of empirical evidence in constructing and testing naturalistic theories of science.

---

[11] Further evidence for Laudan's continued commitment to this form of naturalism can also be found in his later work. See Laudan (1987, 1996).

### 4.1 The use of empirical evidence in theory construction

It now seems obligatory to begin any discussion of the role played by empirical evidence in the philosophy of science with the famous opening sentence of Kuhn's *The Structure of Scientific Revolutions*:

> History, if viewed as a repository for more than anecdote or chronology, could produce a decisive transformation in the image of science by which we are now possessed. (Kuhn 1970, p.1)

Although naturalized philosophers of science generally reject the "image of science" Kuhn reached from his historical studies they continue to acknowledge his role in bringing to our attention the importance of history in developing a descriptively accurate theory of science. Indeed, we can view naturalized philosophy of science as the latest instantiation of a historical trend that has seen history of science playing an increasingly important role in the philosophy of science. As I now hope to show this reliance on history is particularly noticeable in the way naturalized accounts of science are constructed.

A common way to think of theory or hypothesis formation in science is to view it as a response to particularly puzzling natural phenomena that are deemed to require some kind of explanation.[12] The important issue here is not whether this concept is an accurate way to think about theory formation in science but rather how it helps to understand the construction of naturalized accounts of science. I want to argue that the construction of the majority of naturalized accounts considered in the last section can be understood in terms of an abductive strategy of theory formation because they all attempt to formulate general hypotheses concerning science based on case studies that are deemed to be representative of science as a whole.

However, as we shall see, the case studies used to construct naturalized accounts of science are not generally taken from the list of historical examples that have traditionally preoccupied philosophers of science. Naturalized Philosophy of Science (NPS) attempts to break away from this tradition in at least two ways. Firstly, many naturalized accounts of science are based on case studies taken from the life sciences. As such they break with a tradition in philosophy of science that has been dominated by examples taken from the history of physics. Secondly, many naturalized accounts

[12] Readers familiar with the history of philosophy of science will recognise this as Peirce's theory of abduction

choose to focus on contemporary examples of scientific theorizing and practice. As a consequence they place less emphasis on providing an account of more distant episodes in the history of science. In what follows I want to show how these two features are reflected in the selection of representative case studies for constructing various naturalized accounts of science.

Perhaps the most orthodox attempt to construct a naturalistic theory of science based on a particular scientific case study is to be found in Kitcher (1993). Although Kitcher's account is littered with references to a variety of episodes in the history of science he makes it quite clear from the outset that one particular case study is particularly important to the development of his account:

> I shall offer an extended illustrative example, on which subsequent discussions will be able to draw. I shall examine some facets of the history of evolutionary ideas, from the nineteenth century to the present, with a view to fixing ideas about goals, methods, progress, rationality, individual scientific behaviour, and the social structure of science. (Kitcher 1993, p.10)

Here it seems we can see Kitcher advocating something like an abductive strategy to theory formation. Kitcher finds certain features of science interesting but does not yet know how we might be able to explain them. He then identifies a case that is deemed to be representative of scientific development in the hope that we may be able to draw some general conclusions about the puzzling features he is interested in. If this example is genuinely representative then these conclusions may be usefully applied to other areas of science. For example if we can show that the history of evolutionary ideas exhibits a kind of theory change that is both rational and progressiveness then perhaps we can go on to prove that this is the case for the history of other scientific disciplines. For Kitcher this reliance on a detailed case study is not merely heuristically useful but in fact a necessary precondition of constructing a naturalistic account of science:

> It seems to me to be impossible to elaborate a convincing picture [of science] without providing one, fairly detailed, illustration. (Kitcher 1993, p.9)

Further details of Kitcher's conclusions concerning the features he is interested in and the case study he bases them on will be considered in following sections. I now want to move on to another account that is based on a study of the history of evolutionary ideas - the evolutionary account of science of Hull (1988) introduced in the previous section.

Like Kitcher, Hull (1988) chooses to base his naturalized account of science on a detailed example taken from the history of biology. However, Hull is concerned not only with the history of evolutionary ideas but also with the history of systematics from Aristotle to the present day. At the outset Hull makes it quite clear why a detailed discussion of these examples is necessary:

> I intend to use at least *some* of my illustrations as evidence, and the requirements for evidence are a good deal higher than for illustrations. (Hull 1988, p.18)

In fact, Hull believes that the majority of historical examples lack the necessary detail or are too vague to meet these requirements. He even believes that this is true of his own discussion:

> I follow the interconnected paths of evolutionary biology and systematics, starting with Darwin for evolutionary biology and Aristotle for systematics. I cover the early history of these two disciplines quite briefly, too briefly for them to count as evidence. As I approach the present, the narrative slows down appreciably and becomes more substantive. (Hull 1998, p.20)

The latter half of this quote shows that it is Hull's discussion of the more recent features of his historical discussion that function as the real evidence for his account of science. However, the more distant features of these historical narratives still serve an important function for Hull because they enable him to outline his claims about the social structure of science.[13] In this respect at least Hull's approach mirrors that of Kitcher in the sense that they both use historical examples as a resource for illustrating their claims about science.

A similar insistence on the importance of examples taken from contemporary scientific practice is found in Giere (1988). As we saw in the previous section Giere is committed to developing a naturalized account of science that is based on a variety of theories taken from the cognitive sciences. Given this perspective one might suppose that in discussing the nature of scientific theories it would be sufficient for Giere to begin by importing certain cognitive theories that provide an understanding of human cognitive capabilities, particularly those concern the notion of representation. However, Giere makes it clear that we must not assume from the outset that cognitive theories developed to explain the character of everyday representations will be easily applicable to the special case of representations in science, i.e. scientific theories. In other words, if Giere's cognitive approach to the nature of scientific theories is to be efficacious then he needs to show that the representational character of scientific theories is sufficiently similar to the character of other more familiar human representations. As Giere notes, the only way to do this is through a consideration of the structure and function of actual scientific theories:

> I shall not, however, approach the question of what scientific theories might be by looking first at what cognitive scientists have to say about how, in general, humans represent their world. I shall begin, instead, with *scientific* representations themselves, keeping in mind that scientists, after all, are only human. The representations scientists construct cannot be too radically different in nature from those employed by humans in general. (Giere 1988, p.62)

As we have already seen in the cases of Kitcher and Hull, this commitment to basing a general account of science on criteria of descriptive adequacy raises two important issues concerning the selection and treatment of one's preferred examples. In the context of Giere's approach to the nature of scientific theories these general concerns translate into two specific questions. Firstly, how does Giere propose to select an example (or examples) of a scientific theory that is representative in a way that will allow him to draw general conclusions about scientific theories from it? Secondly, given that we can identify a genuinely representative example of a scientific theory, what is the best method of analysis for teasing out its important features?

Let us begin then with Giere's response to the first question, which concerns the selection of a suitably representative example of a mature scientific theory. The first thing to note here concerns the kind of example Giere is after. For unlike Kitcher and Hull, who both focus on the diachronic development of a particular kind of scientific theory, Giere prefers a more synchronic approach:

> The task here is not to reconstruct the historical development of any science but simply to describe, in general terms, the character of a theory as it is understood by contemporary scientists. (Giere 1988, p.63)

Giere also differs from both Kitcher and Hull in that he chooses to base his account on theories of classical mechanics rather than the theory of evolution. In terms of justifying

---

[13] They also serve as an introduction to some of the evolutionary concepts that Hull employs later in the book (see Hull 1988, p.20)

his seemingly unfashionable focus on an example from physics Giere offers three mitigating facts in his defence:

> A number of facts make it a good place to begin investigating the nature of scientific theories. One is that classical mechanics has been discussed by almost everyone who has written on the nature of theories...Second, classical mechanics is typical of a wide range of theories, particularly, though not exclusively, in the physical sciences...A third reason for beginning with classical mechanics is that the education of many scientists still include physics, and the serious study of physics begins with classical mechanics. (Giere 1988, pp.63-4)

Perhaps the most important of these facts for understanding Giere's focus on classical mechanics is the first because it shows his concern with addressing previous work on the structure of scientific theories. Throughout the course of his discussion he makes it quite clear that the "everyone" in question here refers to the received view of scientific theories primarily associated with the work of the logical positivists. By focussing on classical mechanics Giere hopes to show that even the logical positivists most cherished example cannot be made to fit their account of theory structure.

This leads us nicely on to the concerns raised by the second question, the problem of identifying an appropriate method of analysis for dealing with ones preferred examples. For in Giere's view the inadequacy of the logical positivists account of theory structure results from a failure to attend to the way scientists themselves learn and understand theories. Giere relates this failure to an overall problem in approach:

> Philosophers pursuing what they call foundational studies *begin* with a conception of what a theory ought to look like and seek to *reconstruct* theories in that mold. The question whether their conception matches that actually employed in science is begged from the start. (Giere 1988, p.63)

So for Giere the only way to avoid a question begging account of theory structure is to begin by looking at the actual character of scientific theories. Following Kuhn, he suggests that the best way to do this is to analyse the textbook treatment of scientific theories:

> If we wish to learn what a theory is from the standpoint of scientists who use that theory, one way to proceed is by examining the textbooks from which they learned most of what they know about that theory. (Giere 1988, p.63)

The conclusions Giere draws from this approach will be introduced in the next chapter. Here I want to suggest that the analysis of the accounts outlined above enable us to formulate some general conclusions about the use of empirical evidence in the construction of naturalized accounts of science.

The first thing to say about the accounts discussed in this section is that they all share a commitment to using empirical evidence as a basis for theorizing about science. Of course, this again undermines Stump's tacit assumption that we can distinguish these accounts from Laudan's on the grounds that the latter is exclusively concerned with the use of empirical evidence, particularly from the history of science. In fact, I want to argue that the reliance of these accounts on particular case studies actually improves on Laudan's naturalistic methodology. For although it may be true that Laudan uses historical examples in this way the problem is that he never makes it clear that this use of empirical evidence is distinct from its use as a test of his account of science. The virtue of the accounts of Kitcher, Giere and Hull is that they are always explicit about which of these roles a particular empirical example is supposed to play. This section has shown that the first of these functions involves the use of an empirical case study to provide a point of departure for naturalistic theorizing about science.

However, we have also seen that there is much disagreement over what kind of case study might be appropriate for this purpose. This variety is a result of differences in the selection of 1) a representative case study or example, and 2) an appropriate method of analysis. In terms of 1) it seems to me that we can make three useful distinctions with respect to the accounts discussed above:

1. Physics vs. Biology
2. Synchronic vs. Diachronic
3. Historical vs. Contemporary

The value of these distinctions is that they enable us to move beyond the kind of simplistic classification that results from Stump's attempt to reduce the character of various naturalized accounts to their use of results from a particular scientific field. For example, according to Stump's reductionist approach the accounts of Giere and Kitcher will be forced into the same classificatory box because they both choose to appeal to various theories in cognitive science. Likewise Hull's account will be sharply distinguished from these 'cognitive' approaches because of its evolutionary character.

Using the three distinctions outlined above we should now be able to see why this kind of reductionist picture is severely misleading because on the one hand it fails to capture the important differences between accounts that it deems to be similar and, on the other, it fails to capture important similarities between accounts that it deems to be different.

Thus in terms of the first distinction it is obvious that Giere's use of a case study from classical mechanics is in stark contrast to the evolutionary case studies of both Kitcher and Hull. Similarly, the second distinction reflects the fact that whereas Giere is interested in the state of science at a particular time, Kitcher and Hull are more concerned with case studies that reveal the historical development or conceptual lineage of particular scientific ideas. In other words Giere's treatment of his chosen example is synchronic, while those of Kitcher and Hull are diachronic. However, although the first two distinctions seem to pair off Kitcher and Hull against Giere this is certainly not the case with the third distinction. For here we can see that there are important similarities between the work of Giere and Hull that contrasts with that of Kitcher. For although both Kitcher and Hull discuss the development of evolutionary theory their respective treatments reflect a fundamental difference in focus. Broadly speaking Kitcher is more interested in historically distant features of this development (e.g. the development of Darwin's theory) whereas Hull seems to put more evidential weight on more recent episodes (e.g. the debate between the cladists and the pheneticists). In this respect, at least, I believe that this feature of Hull's discussion puts him a lot closer to Giere rather than Kitcher because, like Giere, he is concerned with examples that are taken from contemporary scientific practice.

This last point leads us nicely on to the second feature that characterizes the use of empirical evidence in theory construction – the selection of an appropriate method of analysis. Roughly speaking the distinction between historical and contemporary examples is reflected in the selection of a method of analysis that is best suited to these respective cases. Thus Kitcher's method of treating his preferred example is roughly that of an old style historian of ideas. In this respect his approach is not dissimilar to that of Laudan who, as we saw in the last section, is concerned with using results from the history of science to underpin his naturalized account of science. In contrast, the focus of Hull and Giere on contemporary examples of scientific practice is reflected in their use of more recently developed methods of analysis. Thus, Giere suggests that the issue of theory structure in classical mechanics is best approached through a consideration of the textbook treatment of this theory. Similarly, Hull attempts to bring out the important features of modern debates in taxonomic theory by conducting a citation analysis of this field.

In conclusion, we have seen that empirical evidence plays a crucial role in the construction of naturalized accounts of science. However, it has also become obvious throughout the course of this section that there is a great deal of variation in the way this role is interpreted. In the next section I want to show that there is a second sense in which empirical evidence functions in naturalized accounts of science. For all of the accounts discussed here share a commitment to the view that empirical evidence can be used as a test of their substantive conclusions about science. Here, perhaps unsurprisingly, we shall see that the kinds of factors that influence the selection and treatment of case studies at the level of construction similarly effect the selection and treatment of putative test cases.

## 4.2 Testing and the role of empirical evidence

In section 3.1 we saw that perhaps the most important naturalistic feature of Laudan's developing views was his commitment to the use of certain episodes in the history of science as a test of his substantive claims about science. In fact, such is Laudan's belief in the importance of testing that he has even attempted to extend this notion beyond the confines of his own work. Thus, in *Scrutinizing Science*, Laudan et al. (1992) argue for the crucial role historical evidence can play in the evaluation of conflicting theories of scientific change:

> Scientists understand that it is fundamentally important to test theoretical claims against empirical evidence. *Scrutinizing Science* addresses diverse and frequently conflicting claims about how science changes. It seeks to test these claims against well-researched historical cases. (Laudan, Laudan and Donovan 1992, p.xii)

The theories of scientific change in question here are, of course, the usual suspects of the post-positivist period, namely Kuhn, Feyerabend, Lakatos, Toulmin and even Laudan himself. The complaint of Laudan et al. is that although the proponents of these theories argue for the importance of accommodating historical evidence:

> The sad truth is that most theories of scientific change – including those which have won widespread acceptance – have not yet been extensively or systematically tested against the empirical record. The historical examples liberally scattered throughout the writings of these theorists are *illustrative* rather than *probative* and none is developed in sufficient detail to determine whether the analysis fits the case in hand. (Laudan, Laudan and Donovan 1992, p.5)

The purpose of the research programme outlined in *Scrutinizing Science* is to remedy this situation by providing such probative examples.

However, in undertaking this challenge Laudan et al. acknowledge both that the testing of scientific theories of change against historical evidence is no easy matter; and that their exclusive concern with historical case studies as the appropriate test of such theories constitutes a highly contentious methodological move. Because the first of these problems will be discussed at length in the next section I will gloss over it here. More important for the purposes of this section is the second problem because it alludes to the fact that, in concentrating on the historical case study, Laudan *et al* choose to ignore a variety of other methodological options. These include, but are not exhausted by, the use of questionnaires, statistical analysis, experiments, surveys and ethnomethodological studies.

Although this summary of Laudan's extra-curricular naturalism is incomplete in many respects I believe that it highlights two important features of the role of testing in other naturalized accounts of science. Firstly, all the naturalized accounts discussed in this section take very seriously the need to test philosophical theories of science against empirical evidence. In this sense they share Laudan's complaint that past theories have failed to live up to this methodological maxim. Secondly, although a commitment to testing is common to the accounts discussed in this section, not all of them rely on historical case studies as the appropriate test of their claims about science. As we shall see this is particularly evident in the cases of Giere and Hull who, for different reasons, employ precisely those kinds of methods that Laudan et al choose to ignore. However, before discussing these alternative approaches I now want to show that at least one of the accounts under consideration here follows Laudan in judging historical case studies to be the appropriate unit of testing in naturalized philosophy of science, namely Kitcher (1993).

Given the discussion section 3.1 it is perhaps unsurprising that it is Kitcher's approach to testing that most closely resembles Laudan's preoccupation with historical case studies. For in testing his theory of scientific change Kitcher exhibits a preference for the same kind of historical examples that is evident in his selection of Darwinism as a suitable empirical foundation for the construction of his account. Roughly speaking, Kitcher's account of scientific change is based on the claim that progress in science comes in two fundamental forms: conceptual and explanatory. The details of these two forms of progress will be discussed in the next section, more important here is the way in which Kitcher attempts to establish their descriptive adequacy. For, like Laudan, Kitcher makes it quite clear that the only way to do this is to test his account of progress against the historical record. Thus, in the case of conceptual progress, Kitcher attempts to show that his account is preferable to Kuhn's on the grounds that it provides a better explanation of the shift from the phlogiston theory of Priestley to the oxygen theory of Lavoisier. The problem with Kuhn's account according to Kitcher is that it fails to capture many important features of this historical episode:

> Divided by the chemical revolution, Priestley and Lavoisier inhabit different worlds (or, to put it another way, their theories have different ontologies). Unfortunately, this strategy of accounting for Priestley's apparent success will only accommodate part of the historical evidence. As we insist on the "many worlds" approach, the revolutionary divide between Priestley and Lavoisier looks ever more difficult to bridge, and communication between the two men appears partial at best. (Kitcher 1993, pp.97-8)

Kitcher's naturalistic claim is that his own account of conceptual progress is far better equipped to explain these puzzling features.

A similar concern with the historical record is also evident in Kitcher's discussion of his second argument for scientific progress. Here he attempts to show that his account of explanatory progress can explain the historical development of explanations in atomic chemistry:

> Dalton proposed to explain facts about the course of chemical reactions (and about the weights of reactants and products) by appealing to premises about atoms, premises that specify the "fixed proportions" in which atoms of different elements combine. Chemists ever since have endorsed the claim that this is a correct picture of the objective dependencies, and they have further articulated Dalton's schema. (Kitcher 1993, p.106)

This example, along with that provided for the case of conceptual progress, shows that Kitcher, in terms of testing his key claims about progress in science, chooses to employ a methodology that is based on the use of historical case studies.

Although, as we shall see in later sections, Kitcher is keen to distance himself from many of Laudan's claims it seems that, at least in this respect, he shares with Laudan the following kind of view:

> Theories of scientific change deal with long-term as well as short-term shifts. Guiding assumptions often persist for centuries, revolutions more typically take decades (the Copernican revolution, the Newtonian revolution) than years. History alone can give us any access to those. (Laudan, Laudan and Donovan 1992, p.12)

Of course this similarity between Laudan and Kitcher is hardly surprising given that they are both concerned with developing theories of scientific change. For, as Laudan et al. note, a focus on contemporary examples of scientific change is unlikely to provide the kind of long-term perspective this project requires. However, the commitment to using historical case studies shared by Laudan and Kitcher is far from trivial for, as I will now show, other naturalized accounts of science can and do argue for the use of alternative methods of testing for their preferred claims about the nature of science.

Such an alternative method can be found in Hull's approach to testing his evolutionary account of science. For, as suggested in the previous section, although Hull uses historical examples to illustrate his claims about the social structure of science, in terms of testing these claims he places greater evidential weight on contemporary examples of scientific practice. In fact, Hull can be seen here to be explicitly rejecting the kind of historical approach to testing advocated by Kitcher and Laudan:

> When scientists first opt one way or the other on important issues, the causal circumstances that are relevant to these decisions are extremely particularized. Only an intensive and extensive investigation of these circumstances can explain why science took the course it did. Because these causal situations are so particularized and the requirements for evidence so stringent, rarely can episodes in the history of science serve as evidence for or against particular views about the nature of science. The necessary evidence too often is missing. (Hull 1988, p.21)

In order to avoid this kind of problem Hull suggests that the appropriate test of his account is to show how it can capture important features of contemporary disputes in biology for:

> Like it or not, the sorts of inquiries that hold out the greatest hope of distinguishing between alternative about the nature of science are those concerned with present-day science. Only in such circumstances are the relevant data available. That is why I turned to recent disputes in biological systematics, in particular the fates of phenetics (or numerical taxonomy) and cladistics (or phylogenetic systematics). (Hull 1988, p.21)

For Hull the 'relevant data' referred to here is gathered through the use of techniques usually associated with contemporary sociology, particularly citation analysis and the use of questionnaires.

A similar concern with using more contemporary examples of scientific practice to test naturalized accounts of science is also found in Giere (1988). However, in contrast to Hull (1988), Giere chooses to focus on the laboratory study as the appropriate way of testing his naturalized account of science:

> Advocates of laboratory studies regard their observations in the laboratory as providing empirical support for a constructivist interpretation of science. I propose to counter their claim with a laboratory study of my own. My subjects are nuclear physicists working in a national cyclotron facility. These scientists regard themselves as investigating the structure of the nucleus by bombarding various nuclei with rapidly moving light nuclei, mainly protons, and seeing what comes out. (Giere 1988, p.111)

Giere's aim here is to show that, contra to the anti-realist conclusions of the constructivists; at least one example of laboratory practice is best explained by his own realist account of scientific practice. However, in advocating this approach, Giere is keen to point out the restricted nature of his conclusions:

> This finding, of course, will not mean that empiricism or constructivism might not apply in some other areas of science. Which type of account applies must be decided on a case-by-case basis. Only then might one seek some rough generalizations regarding the kinds of circumstances in which realism, rather than empiricism or constructivism, provides. (Giere 1988, p.112)

A clearer statement of the naturalist's commitment to the empirical testing of philosophical theories of science could not be asked for. In fact, I want to conclude my discussion of this feature of naturalized philosophy of science by suggesting that this statement captures the methodology at the heart of all the accounts discussed in this section. For although we have seen that there are important differences between these accounts, both in the claims they make and in the nature of cases that are used to test them, they all rely on the notion that the possibility of establishing normative claims concerning science is crucially dependent on the provision of supporting empirical evidence.

Before moving on to a discussion of the role that these normative claims play in naturalized accounts of science I want to return to an issue that was left open at the end of the last section. There I argued that Stump's reductionist concept of a 'metamethod' was unable to provide an adequate account of the complexity of naturalism in the philosophy of science. This conclusion was argued to result from the inability of this approach to capture two important features of the accounts discussed throughout this section. Firstly, contrary to Stumps reductionist claim, naturalized philosophers of science generally rely on more than one science as a resource for constructing their accounts of science. Secondly, and more importantly, in discussing Laudan's approach we saw that his reliance on the history of science as a test of his claims did not seem to be the same as the appeal of other naturalist's to the results of particular sciences. Consequently I suggested that we needed an alternative account of these two puzzling features. Utilizing the analysis of this section we are now in a position to provide such an account.

### 4.3 Weak and strong scientistic naturalism

In the previous section, I chose to begin my discussion with an account of Laudan's attempt to test various theories of scientific change against historical case studies. Coupled with the recognition of the similar role played by such studies in testing his own work, this perhaps explains Stump's belief that Laudan's 'metamethod' *is* the history of science. However, we have seen that exactly the same kind of methodological approach is characteristic of all the accounts discussed in this section. This, of course, undermines Stump's claim that the metamethodology behind these accounts can be reduced to their use of a particular branch of science. For example, if Laudan's use of the history of science counts as a metamethod then so must Kitcher's similar reliance on historical case studies. But Stump's account also implies that, where present, the appeal to scientific findings will also count as a metamethod. Thus, in the case of Kitcher's account, we must conclude that there are two metamethodologies at work; one based on his use of historical case studies, the other on his appeal to findings in cognitive science.

This seems to leave us at something of an impasse because the whole point of Stump's approach is to identify a single metamethod at work in each of the accounts he considers. This uncomfortable conclusion, I believe, can be avoided once we recognize the fact that the use of empirical evidence (whether it is historical or contemporary in nature) is functionally distinct from the appeal to the results of particular branches of science. In fact, pace Stump, I want to argue that the former constitutes the true 'metamethodology' of naturalized accounts of science, with the latter functioning, not as a metamethod, but rather as addition to this methodological core. Based on this claim I argue for the following characterization of naturalized philosophy of science:

**Weak Scientistic Naturalism** - The study of science should mirror the methodological practices of science itself particularly in terms of using empirical evidence as a tool for constructing and testing theoretical claims about science.

**Strong Scientistic Naturalism** - In addition to mirroring the methodological practices of science we must also draw on the findings or results of well-established scientific theories.

Using these definitions we can now see what the relevant differences between Laudan and his fellow naturalists are. Because Laudan does not appeal to the results of particular sciences his approach to naturalism can be classified as weakly scientistic only. In contrast, the presence of this feature in the accounts of Giere, Kitcher and Hull marks off their naturalistic approaches as strongly scientistic. My claim is that this way of classifying naturalized accounts of science does far better justice to the status of modern naturalism in the philosophy of science (in terms of the different roles played by empirical evidence on the one hand and the appeal to particular sciences on the other) than does Stump's concept of a metamethod.

In the next section, I conclude the discussion of this chapter by returning to the question that motivated the analysis of this section; does Kitcher's account of traditional naturalism adequately capture naturalism in the philosophy of science? In the context of my discussion this general question leads to another more specific one; does Kitcher's account capture the kind of distinction I have made between weak and scientistic naturalism? In order to answer this question we need to appreciate that there is more to Kitcher's account of naturalism than was discussed in section 1. As I will now show, Kitcher attempts to capture the naturalism in the philosophy of science by suggesting how modifications in the core position of "traditional naturalism" can lead to a variety of alternative positions.

### 5. Kitcher's account of naturalism in the philosophy of science

As we saw in section 2, Kitcher's account of naturalism is based on the notion that traditional (or normative) naturalism occupies an uneasy position between earlier a priori epistemologies and more radical (or anti-normative) forms of naturalism. Kitcher attempts to cash out this instability by showing how proponents of both these alternatives can interpret particular problems with traditional naturalism as supporting a move toward their own positions. I will now attempt to show that, as with naturalized epistemology, Kitcher takes this kind of account to be equally applicable to naturalized philosophy of science *and* its competitors.

According to Kitcher, the first possible complaint against epistemological naturalism could take the following form:

> Complaint (A) – The usual philosophical sources of normative principles are not displaced by traditional naturalism, which offers only the meta-epistemological principle that the deliverances of these sources are not a priori. (Kitcher 1992, p.78)

In the context of epistemology we saw that this complaint could be used by Post-Fregean epistemologists to show that traditional naturalism leaves untouched the status of their project, or it could be used by the radical naturalist to suggest that traditional naturalism does not go far enough in its rejection of earlier epistemologies. The crucial point here is not whether this is true of epistemology but rather how Kitcher fits the status naturalized philosophy of science into this picture. Is the situation in this field the same as epistemology? As we might expect Kitcher answers this question in the affirmative for he claims that the radical naturalists response to this complaint can be found in the work of the Edinburgh school of the sociology of scientific knowledge.

At this point we can recall that Kitcher suggested that one way the traditional naturalist might respond to this problem would be to draw on the third thesis of traditional naturalism. This thesis, concerning the relevance of empirical information for the formation of normative principles, suggests that complaint (A) is untenable. If we are after a normative account that will promote cognitive success in actual epistemic contexts then empirical information about how we function in such contexts must be strongly relevant to our project. As Kitcher notes this undercuts the force of complaint (A) but creates a further problem concerning the seemingly circular nature of this response:

> If proper epistemic recommendations are crucially dependent on contingent information about the world, how could we acquire the information on which those recommendations depend? (Kitcher 1992, p.79)

This in turn gives rise to the following complaint:

> Complaint (B) – Only if we can arrive at principles that would properly guide inquiry in any world and which can be validated a priori will the problem of normative epistemology be solved. (Kitcher 1992, p.79)

As for complaint (A), Kitcher shows how both Post-Fregean epistemologists and radical naturalists can use this complaint to argue against the normative project of traditional naturalism. For the former such an appeal requires the rejection of theses (3) and (4) of traditional naturalism, for the latter thesis (4) – the rejection of the a priori status of normative principles – coupled with (B) seems to show that traditional naturalism is impossible.

However, in defence of traditional naturalism, Kitcher points out that this kind of argument against it is based on a particularly strong form of Cartesian scepticism:

> The sceptic's demand is for synchronic reconstruction of beliefs: take the totality of things you believe, subtract this claim and everything that you cannot defend without assuming it, and now show that the claim is correct. (Kitcher 1992, p.90)

But, of course, the whole point of the naturalist's position is that they do not believe that such a reconstruction is possible because there is no body of a priori knowledge that we can appeal to carry out this project. Kitcher believes that this shows that the use of complaint (B) to reject traditional naturalism completely misses the point behind its conception for:

> On naturalism's own ground, there are bound to be unanswerable forms of scepticism. Traditional naturalists should therefore decline blanket invitation to play the game of synchronic reconstruction. Each of us absorbs information from our predecessors, and, through our own interactions with nature and with one another, we modify our collective picture of the world and of the proper ways to investigate it. (Kitcher 1992, p.90)

Thus, according to Kitcher, traditional naturalism is unconcerned with traditional forms of scepticism. The importance of this claim for the discussion of this section is that Kitcher suggests that the move to naturalism replaces these traditional concerns with new, and perhaps more difficult, sceptical questions. As I will show in the next section, the variety of responses to these kinds of questions can be used to explain the features of naturalized philosophy of science discussed in sections 3 and 4.

### 5.1 Traditional naturalism and the philosophy of science

Kitcher suggests that the traditional naturalist is committed to the following kind of view:

> Naturalism offers the optimistic picture of a particular type of organism, beginning with rudimentary representations of nature and primitive notions of how to modify those representations, and gradually replacing these with cognitively representations and strategies. (Kitcher 1992, p.90)

His claim is that we can understand modern, as opposed to traditional, sceptical challenges to this picture as revolving around two possibilities:

1. The possibility that we began in so primitive a state that we are incapable of working ourselves into any accurate representation of nature, and;
2. The possibility that there are constraints on the processes of modification that prevents us from making significant improvements.

Accordingly, the naturalist is faced with two ways of responding to our reformed sceptic. She can either provide arguments to show that we could not have begun in the kind of primitive state suggested in (i); or she can attempt to show that our self-corrective mechanisms are, contrary to (ii), powerful enough to guarantee improved representations of nature *no matter how we started.*

As Kitcher notes, the first of these responses to the sceptic is most closely associated with the appeal to evolutionary theory, which results in the following kind of argument:

> If our initial cognitive equipment were as unfortunate as the sceptic portrays it as being, then, the suggestion runs, our ancestors would have been eliminated by natural selection. (Kitcher 1992, p.91)

Kitcher, rightly I believe, regards these kinds of arguments as extremely inconclusive because they seem to provide nothing more than speculation.[14] However, for our

---

[14] In fact, Kitcher's impression is that current evolutionary thinking actually reinforces (i) rather than being an argument against it.

purposes the important point to note here is that in discussing this kind of appeal to evolutionary theory Kitcher is able to show how the EEM programme (discussed in section 3.1.2) relates to more conservative attempts to naturalize the philosophy of science.

Far more important for Kitcher is the second kind of sceptical concern, which questions the belief that we can make significant improvements in our representations of nature. This feature of Kitcher's account is also crucially important to our discussion because it is essentially an attempt to capture the nature of modern debates between normative and anti-normative approaches to naturalism. For, although many proponents of the latter have rejected the former by appealing to traditional sceptical arguments, the really damaging objections in Kitcher's view arise from attempts to show that the kind of self-correction required by traditional naturalism is not reflected in the history of science. As Kitcher notes:

> The root of this form of scepticism lies in arguments about the underdetermination of belief by encounters with nature.

Here Kitcher is drawing attention to the fact that many contemporary sociologists of science have generally attempted to block the conclusions of their cousins in philosophy by appealing to a variety of factors that seem to undermine the latter's normative claims.[15]

For Kitcher, this reformed version of scepticism explains the prevalence of empirical examples in the naturalized accounts discussed in sections 4.1 and 4.2. Here it is worth quoting Kitcher at length:

> If I am right, (ii) is the dangerous form of scepticism, threatening to collapse traditional naturalism into a radical position that abandons or relativizes normative epistemology. The general arguments for that collapse are not cogent, and assessing the viability of traditional naturalism turns on a number of interesting questions: How penetrable is perception by cognition? What kinds of systems of authority inhibit or promote change? How is the significance of a problem or an accomplishment appraised? How are instruments and experimental designs assessed? How do social and cognitive interests combine in scientific decision making? Answers to these questions must be sought in the context of detailed studies of historical and contemporary scientific practice, if we are to determine whether science is an instrument of self-correction (as the traditional naturalists

[15] These include arguments from shifting standards, the theory-ladenness of observation, assessment of experiments, social embedding and the effects of authority (for details see Kitcher 1992, p.94-5)

> would have it) or whether it is simply a vehicle for the expression of different, incommensurable, forms of life. (Kitcher 1992, p.100)

This is Kitcher's way of capturing what I have previously referred to as 'weak scientistic naturalism' – the use of empirical evidence in constructing and testing normative theories of science. However, here Kitcher makes it clear that this evidence is being used to decide disputes between normative and anti-normative approaches to naturalism. This is in contrast to my approach, which focuses on the use of empirical evidence to decide disputes that primarily arise within the normative camp only.

However, having established that the sceptical battleground for contemporary naturalistic approaches rests on the appeal to appropriate forms of empirical evidence, Kitcher attempts to show that a similar appeal is relevant even for those who reject the sceptic's move. Here Kitcher shows that a further argument against the traditional naturalist can be made in terms of the following complaint:

> Complaint (C) – The History of science reveals that the goals attributed to inquiry vary widely from field to field and from epoch to epoch. There can thus be no universal normative epistemology. (Kitcher 1992, p.80)

This kind of conclusion should be familiar to the reader from my discussion of Laudan (section 3.1.4). Thus, in the context of Kitcher's account, we can read Laudan's claims about the transitory nature of goals as saying that the best the traditional naturalist can hope for is to construct normative principles that are relativized to a particular period of inquiry based on its own conception of cognitive goals. Again we can see here that Kitcher allows for another feature of the accounts discussed in the last section. For Kitcher is able to accommodate Laudan's naturalism by suggesting that it constitutes a modification of the traditional naturalists commitment to providing a universal goal for science. In this way, Kitcher's claims that his own, and others (e.g. Giere or Hull), commitment to providing such a goal constitutes the core project of traditional naturalism. Kitcher is also keen to point out that even for those naturalists who accept the goal of constructing a normative account (whether it is universal or local) there are still possibilities for disagreement. These centre on the nature of representation and the role of sociological factors in science. Like Kitcher, I shall deal with these briefly.

The fourth complaint Kitcher claims can be levelled at traditional naturalism is that its treatment of representation is overly propositional. To the extent that much cognitive science emphasizes the non-propositional nature of mental representation this should be reflected in a naturalistic approach. This results in complaint (D):

> Complaint (D) – Traditional epistemology thinks of knowledge as primarily propositional. This presupposition should be scrutinized in the light of historical and sociological analyses. Where necessary, the standard epistemological idioms of belief and justification should be absorbed within a broader vocabulary or discarded entirely. (Kitcher 1992, p.81)

Here we can see Kitcher allowing for approaches like Giere (1988), which reject the propositional approach to representation in favour of the semantic approach to theory structure. Kitcher also suggests that proponents of a normative naturalistic approach to the study of science may object to its emphasis on the cognitive lives of individuals independently of their situation in social contexts. Based on this he offers a further complaint of traditional naturalism:

> Complaint (E) – Epistemology must examine the attainment of knowledge by communities as well as by individuals, and should investigate strategies through which communities could advance their epistemic ends. The appropriate strategies for individuals to follow cannot be identified without considering the communities to which they belong. (Kitcher 1992, p.82)

Again we can see that this modification of traditional naturalism captures sociological approaches such as that provided by Hull (1988) who, as we have seen, argues for the collective rationality of science rather than the individual rationality of particular scientists.

This completes my outline of Kitcher's framework for understanding epistemological naturalism, particularly as it exists in the philosophy of science. We have seen that complaints (A) and (B) serve to distinguish normative naturalized philosophy of science from anti-normative sociology of science. Complaint (C) allows Kitcher to account for an approach like Laudan (1984) who argues for a restricted normative account of science based on the transitory nature of its goals. Complaints (D) and (E) can be seen as a method of differentiating normative naturalistic approaches in terms of the possible replacement of the overly propositional and/or individualistic nature of the traditional naturalism. The resultant picture is of a set of loosely connected naturalistic approaches to the study of science that radiate from the core position of "traditional naturalism."

### 5.2 Assessing traditional naturalism

It seems then that Kitcher's account of epistemological naturalism can capture many of the features that sections 3 and 4 showed to be characteristic of naturalized philosophy of science. For Kitcher, the variety of claims and approaches found in these accounts is best explained in terms of the sceptical idiom of contemporary epistemology. In this way, he is able to show that the core claims of traditional naturalism are an attempt to respond to a variety of sceptical challenges. However, although we may be tempted to conclude that Kitcher (1992) provides the appropriate framework for discussing the status of these accounts, I now want to argue that it is deficient in the way it makes certain naturalized accounts peripheral to the traditional naturalist's project.

An important feature of Kitcher's account of naturalism is the way it attempts to provide a 'snapshot' of the state of naturalism at a particular time. For although Kitcher provides us with an historical account of the development of traditional naturalism, his treatment of contemporary naturalism is based on a synchronic analysis of the relations between traditional naturalism and its competitors. My worries concerning this approach are twofold. Firstly, the historical account Kitcher provides, although relevant to the development of naturalism in the philosophy of science, seems better suited to explaining the development of naturalism in contemporary epistemology. Secondly, and perhaps more importantly, Kitcher's claim that 'traditional naturalism' forms the core position from which other naturalized accounts deviate seems to overplay the role of his own preference for this form of naturalism over others. In fact, these two worries are intimately connected for I believe that Kitcher's preference for a particular form of naturalism leads him to produce an historical account that privileges developments that support this preference. In other words, my charge against Kitcher is that his account of 'traditional naturalism' whiggishly reads back into the history of philosophy his own naturalistic preferences.

Of course, whiggism is not always a bad thing. As Laudan points out, there is often nothing beyond historical judgments of progress other than the fact that they reflect our interest in pursuing certain cognitive goals. However, in the case of Kitcher's account of naturalism, I do not believe we can accept this kind of defence. Ironically, my reasons for this claim come from Kitcher's treatment of Laudan's naturalized account of science (see Kitcher 1993). Recall that, for Kitcher, this account constituted a modification of the traditional naturalist's commitment to providing a universal goal for science. My worry is that in marginalizing Laudan's account in this way, Kitcher has missed an important feature of the development of naturalism in the philosophy of science. For, as I will show in the next chapter, Laudan's arguments concerning the transitory nature of goals in science (and, more importantly, his pessimistic against realism) provided one of the most important motivations behind the development of realist accounts within naturalized philosophy of science. Thus, in contrast to Kitcher's synchronic approach, what we need is an account that captures the important features of this development.

## 6. Conclusion

In this chapter we have attempted to get an idea of the sorts of claims and arguments that are characteristic of recent attempts to naturalise the philosophy of science. We have seen that a shared commitment to naturalism in the philosophy of science is compatible with a variety of opinions concerning just what a suitably naturalized philosophy of science should look like. In particular, significant disagreement was shown to exist between naturalized philosophers of science concerning how we should answer two important questions. Firstly, given that there is no longer any injunction against using scientific theories to address philosophical problems, which (if any) scientific discipline should we turn to? Secondly, given that empirical evidence is relevant to the adjudication and construction of different philosophical theories about science, what kind of evidence do we need to construct such theories and how should we test them? By using these two questions as a general classificatory framework I was able to show some of the important features that connect the work of naturalized philosophers of science like Kitcher, Giere, Hull and Laudan.

# Chapter 2

# The Genesis of Naturalistic Realism

## 1. Introduction

So far we have only discussed the way in which naturalised accounts of science attempt to construct and test various philosophical theses about science. We have not yet discussed what these particular philosophical theses might be. In this chapter I show that most (but not all) attempts to naturalise the philosophy of science aim to defend some form of scientific realism. However, as suggested in the previous chapter, I do not think that Kitcher's account of epistemological naturalism is best placed to explain this particular feature of naturalised philosophy of science. So, in the next section, I introduce Rosenberg's account of philosophical naturalism and show that it better explains the role of realism in naturalised philosophy of science. As a consequence of this analysis I argue that naturalistic realism is primarily a reaction to three particular antirealist accounts of science: Laudan's naturalistic antirealism, van Fraassen's constructive empiricism, and various 'constructivist' accounts provided by recent sociology of science. In sections 3, 4, and 5 I discuss the main features of each of these accounts respectively. In section 6, I introduce three varieties of naturalistic realism and show how each of them is motivated to refute or challenge the arguments and claims of these three antirealist positions.

## 2. Rosenberg's philosophical naturalism

In his 1996 article, "A Field Guide to Recent Species of Naturalism", Alexander Rosenberg is concerned with understanding naturalism as a form of philosophical inquiry employed by both realists and antirealists in contemporary philosophy of science. Although Rosenberg shares Kitcher's emphasis on epistemological issues as the main factor behind the recent resurgence of naturalism, he disagrees over the details.[1] This is reflected in his account of 'philosophical naturalism,' which he suggests is committed to the following three "axioms":

[1] For example, although Rosenberg considers the work of Quine (1951) as crucial in clearing the ground for naturalistic approaches within philosophy, he suggests that it was Nagel (1961) who first advocated the use of scientific theories in addressing philosophical problems. He bases this claim on the fact that

1. The repudiation of 'first philosophy'. Epistemology is not to be treated as a propaedeutic to the acquisition of further knowledge.

2. Scientism. The sciences - from physics to psychology and even occasionally sociology, their methods and findings - are to be the guide to epistemology and metaphysics. But the more well established the finding and method the greater the reliance philosophy might place upon it. And physics embodies the most well established methods and findings.

3. Darwinism. To a large extent Darwinian theory is to be both the model of scientific theorizing and the guide to philosophical theory because it maximally combines relevance to human affairs and well foundedness.

(Rosenberg 1996, p.4)

With the exception of axiom 1 (which is almost identical to thesis 4 of Kitcher's traditional naturalism), it is immediately obvious that Rosenberg's understanding of naturalism differs from that of Kitcher (1992), even if this is simply a disagreement over the appropriate idiom in which to discuss the issues.[2] Before attempting to assess the value of Rosenberg's account it might be useful to compare philosophical naturalism with Kitcher's traditional naturalism.

## 2.1 Philosophical naturalism vs. traditional naturalism

At first glance, none of Kitcher's four theses of traditional naturalism seem to correspond to Rosenberg's second axiom concerning the use of science in epistemology and metaphysics. However, on closer inspection, the most suitable candidate would seem to be Kitcher's second thesis:

> (2) The epistemic status of a state is dependent on the processes that generate and sustain it. (Kitcher 1992, p.75)

This thesis can be read as a veiled reference to the scientism emphasized by Rosenberg because in maintaining the importance of "processes that generate and sustain" Kitcher is implicitly suggesting the use of psychology and/or sociology (i.e. the sciences that deal with such processes) in epistemology (i.e. the subject that deals with "the epistemic status of a state"). However, despite this apparent agreement on the scientistic basis of

---

while many naturalists regard Quine's rejection of traditional empiricism as sound they have not been prepared to accept many of Quine's more radical conclusions. See Rosenberg 1996, pp.1-2.

naturalism, there are two important points that need mentioning here. Firstly, Rosenberg's second axiom is more general than Kitcher's second thesis in the sense that it is intended to capture both epistemological *and* metaphysical naturalism. Secondly, in emphasizing the sheer variety of sciences that are employed as a source of both methods and findings, Rosenberg's second axiom seems to be a better way of capturing the scientism that exists in naturalized philosophy of science. *Pace* Kitcher, epistemological naturalism is not simply the use of a particular science (psychology) in a particular branch of philosophy (epistemology).

Rosenberg's third axiom concerning the role of Darwinism is entirely absent from Kitcher's formulation of "traditional naturalism." Kitcher does discuss Darwinism but he does not consider it important enough to include in his core definition. Rosenberg insists that the influence of Darwinism on contemporary naturalism is widespread and important enough to grant it special status.[3] When one looks at fields like evolutionary ethics and evolutionary epistemology Rosenberg certainly seems to have a point. However, one could argue against the inclusion of Darwinism as an "axiom" in Rosenberg's general framework given that it is already covered by the more general scientistic axiom 2. I do not wish to go into detail on this point but merely want to add the caveat that there are some naturalistic approaches that do not rely heavily on Darwinian theory. In fact, as was demonstrated in chapter 1, even when philosophers do invoke Darwinism they do not all do so with the same emphasis. Having said this even if Rosenberg is wrong to include Darwinism as an "axiom" of philosophical naturalism we can still gain a valuable understanding of most naturalistic approaches from his first two axioms.

Despite their different interpretations of the best way to characterize epistemological naturalism, Kitcher and Rosenberg are at least in agreement in saying that it is generally put forward as a compromise (or mediator) between competing positions:

> Traditional naturalists occupy an uncomfortable middle ground between earlier epistemologists and those who campaign for abandoning (or relativizing) normative projects. (Kitcher 1992, p.77)

---

[2] The very language Rosenberg employs (in contrast to Kitcher's use of idiomatic phrases from traditional epistemology) suggests that he is approaching his subject from quite a different perspective.

[3] For Rosenberg this special status has its origin in two features of Darwinism. The first concerns the teleological nature of evolutionary theory. The second is the increasingly widespread acceptance of evolutionary theory as an explanatory device for understanding existing human traits.

> Naturalists in most of the subdisciplines of philosophy have viewed their project as one that seeks a reasonable compromise between competing extremes while doing as much justice to the extreme views as they deserve. (Rosenberg 1996, p.2)

This is best seen as a shared commitment to a general framework (Figure 1):

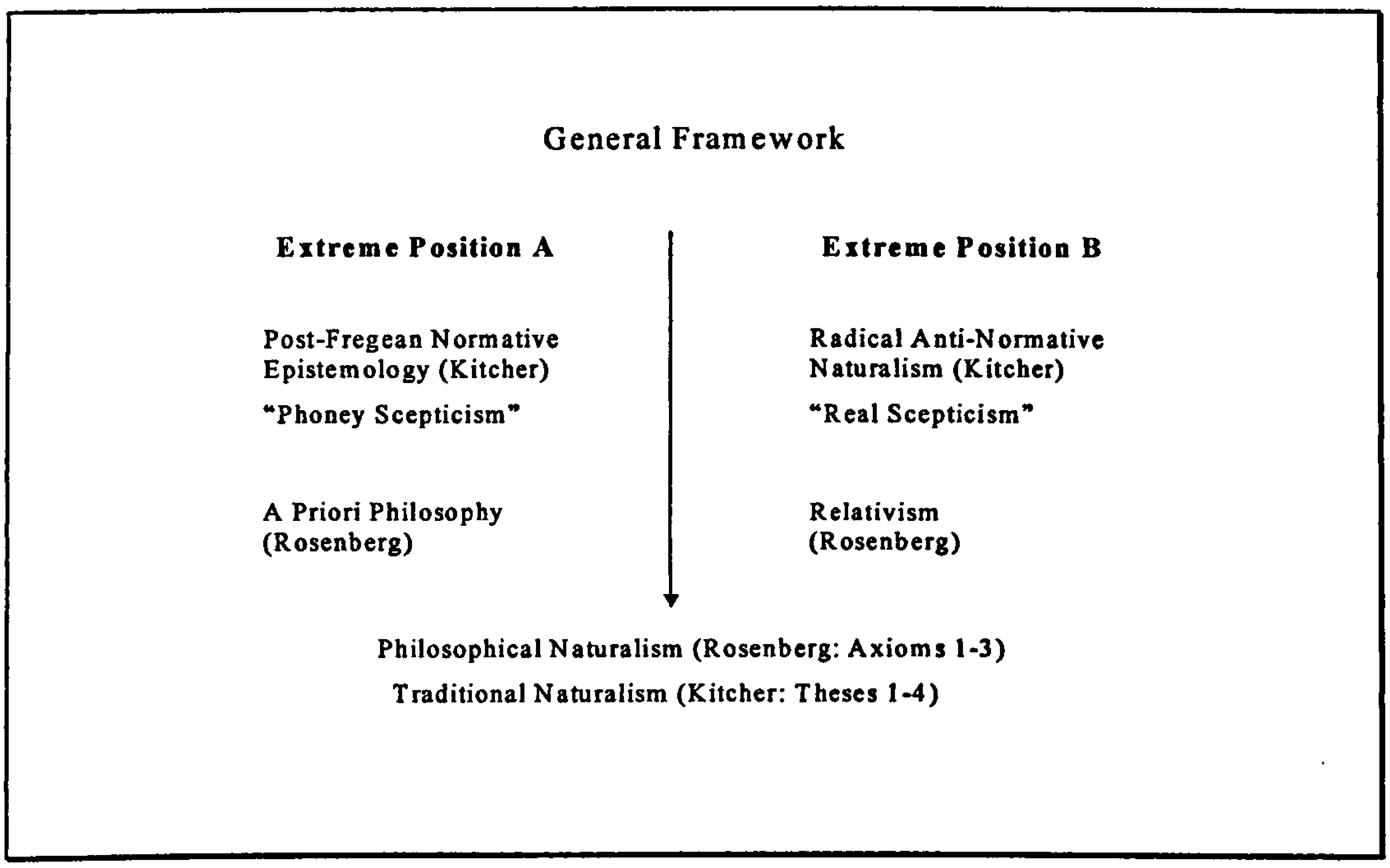


**Figure 1**

Although the framework shown in figure 1 clearly captures an important structural similarity between the accounts of Kitcher and Rosenberg, there are also obvious differences. Firstly, the "extreme" positions that Kitcher and Rosenberg put forward again illustrate the contrasting nature of their respective projects. Thus, Rosenberg goes for the more general dispute between a priori philosophy and relativism whereas Kitcher settles on the narrower, specifically epistemological, dispute between Post-Fregean normative epistemology and radical anti-normative naturalism. These contrasting interpretations of the general framework again illustrate the more general character of Rosenberg's philosophical naturalism as opposed to the narrow epistemological focus of Kitcher's traditional naturalism.

The key issue that separates Kitcher from Rosenberg is *normativity*. It is clear from his characterisation of competing extremes that Kitcher takes disagreements over the normative status of epistemology to be the driving force behind traditional naturalism. The purpose of traditional naturalism is to preserve this status (thesis 1) without thereby accepting the apsychologism (thesis 2) and apriorism (thesis 4) that are characteristic of traditional normative epistemology. Now when we consider Rosenberg's characterisation of competing extremes we find no mention of normativity. Indeed, this is reflected in Rosenberg's characterisation of philosophical naturalism for although it contains claims concerning a commitment to scientism (axiom 2) and the rejection of apriorism (axiom 1), there is nothing that corresponds to Kitcher's nomativity thesis. I will now attempt to show how this difference between the general accounts of Kitcher and Rosenberg affect their specific accounts of naturalism in the philosophy of science.

## 2.2 Philosophical naturalism and the philosophy of science

In order to find anything that looks like Kitcher's normativity thesis in Rosenberg's account we must turn to his account of naturalized philosophy of science (NPS). Here Rosenberg suggests that NPS can only be adequately captured by adding the following "theorem" to the three core axioms:

> 4. Progressivity. Arguments from the history or sociology of science to the non-rationality, or non-cumulativity, or non-progressive character of science, are all either unsound and/or invalid. (Rosenberg 1996, p.4)

So, for Rosenberg, a naturalized philosopher of science is someone who accepts the three axioms of philosophical naturalism and *in addition* believes in the rationality and progressiveness of science. This claim is clearly very similar to Kitcher's normativity thesis but it is important not to confuse the two. Although arguments for the rationality and progressiveness of science are often characteristic of normative accounts of science, a normative approach is neither necessary nor sufficient for establishing such concepts as genuine features of science.[4] In other words, a commitment to progressivity and rationality is often accompanied by, but is not the same as, a commitment to normativity.

This difference in focus is further illustrated by Rosenberg's characterisation of the specific motivation behind naturalism in the philosophy of science:

> Naturalism [in the philosophy of science] seeks to reconcile the history of science as a human institution with the positivist insight that its methods are uniquely suited to providing objective knowledge certifiable independent of its social and psychological context. (Rosenberg 1996, p.3)

---

[4] For example, through a purely descriptive historical investigation one could establish that science is, as a matter of fact, progressive without saying anything about whether or not it *should* be.

Again, although this claim may seem to be very close to Kitcher's account, it is not quite the same. For whereas Kitcher discusses the dispute between positivists and radical sociological relativists in terms of disagreements over normativity, it seems that Rosenberg is more concerned with their respective views on the possibility of attaining *objective knowledge*. That this is indeed the case is made clear in the following passage:

> Naturalism [in the philosophy of science] hopes to provide a viable alternative to the constructivist relativism which swept through the social studies of science in the 1980's and the 'constructivist empiricism' (van Fraassen [1980]) of philosophers who despair of a less radical defence of the objectivity of science. (Rosenberg 1996, p.3)

Thus, for Rosenberg, naturalists in the philosophy of science defend progress and rationality in science because they are ultimately concerned with a suitably robust account of the objectivity of scientific knowledge not, *pace* Kitcher, because they are concerned with defending the normative status of their discipline.

In Figure 2 I have attempted to make clear the discussion of the last paragraph by showing how Kitcher and Rosenberg apply the general framework of figure 1 to naturalism in the philosophy of science.

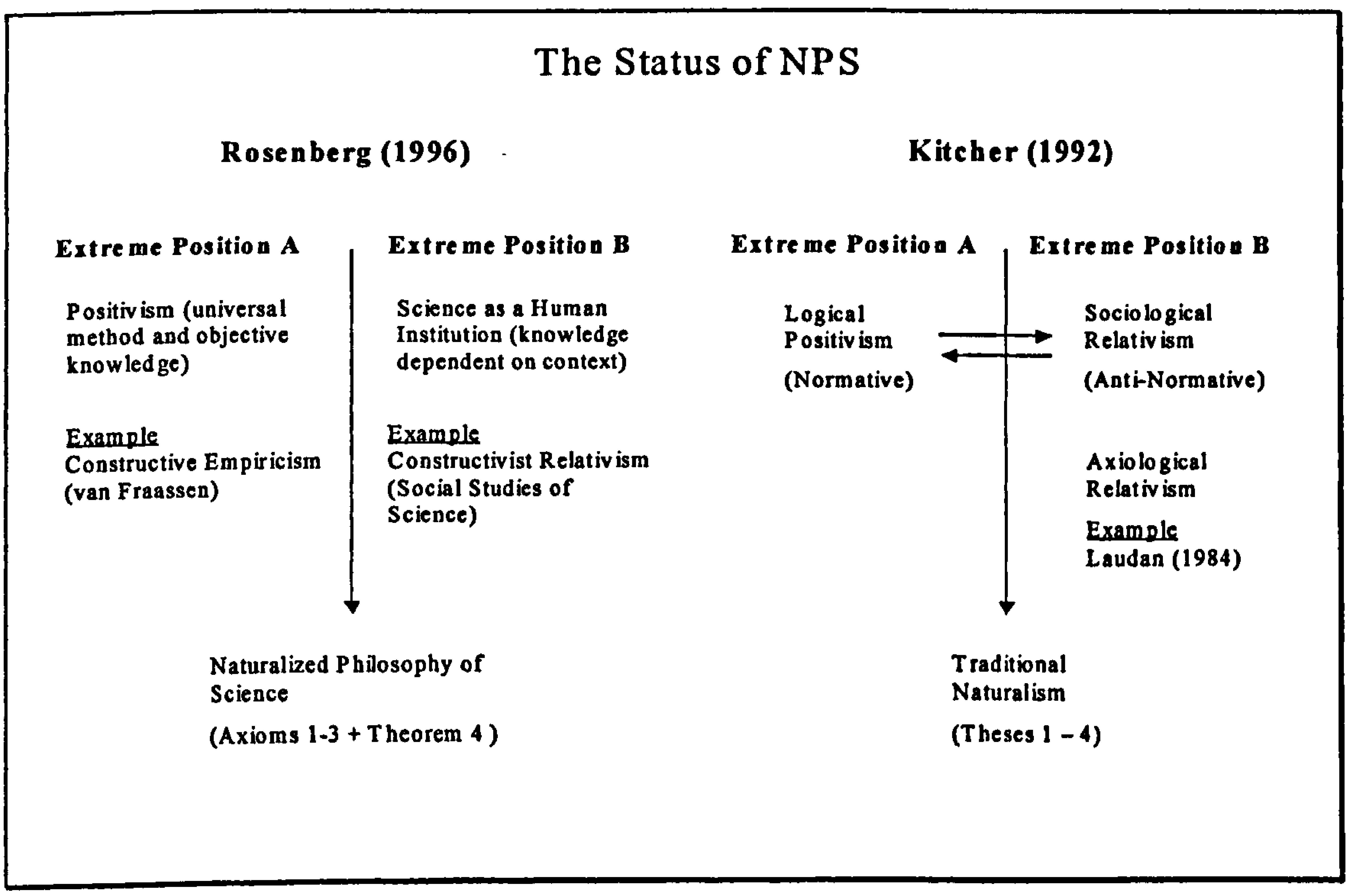


**Figure 2**

Here we can see why it would be all too easy to take Kitcher and Rosenberg as saying essentially the same thing. For although they have different ways of characterizing the relevant issue that splits the old-school positivist from the relativist (objectivity vs. normativity), they at least concur in thinking that it is a dispute between these two positions that drives naturalism in the philosophy of science. However, figure 2 also gives us reason for thinking that this disagreement over how to characterize the positivism vs. relativism dispute may not be as trivial as it appears. Kitcher and Rosenberg agree that, at least in the context of naturalized philosophy of science, the most appropriate candidate for extreme position B is sociological or constructivist relativism.[5] However, in contrast to Rosenberg, Kitcher argues that another opponent of naturalism, at least on the "radical" side of the divide, is axiological relativism. Given its importance to the rest of this chapter I shall conclude this section with a discussion of this important feature of Kitcher's account.

### 2.3 Axiological relativism

In chapter 1 we saw that one of the major complaints levelled at Kitcher's traditional naturalism was that the history of science seemed to rule out the possibility of developing a universal account of cognitive virtue:

> Complaint (C) -The history of science reveals that the goals attributed to inquiry vary widely from field to field and from epoch to epoch. There can thus be no universal normative epistemology, and we must settle for description of the ways in which people actually form their beliefs or for local recommendations about how those working within a particular context should operate to advance their goals. (Kitcher 1992, p.80)

Kitcher makes it quite clear that, if this complaint is correct, then there is no hope for the normative epistemology (or for that matter philosophy of science) because universality is constitutive of this project. In Kitcher's view, epistemology is nothing if it cannot provide an objective standard for assessing epistemological and methodological principles.[6]

---

[5] This is because relativism of this variety tends to satisfy both Kitcher's major criterion of being anti-normative and Rosenberg's of being opposed to the development of an objective account of scientific knowledge (at least insofar as philosophers conceive of this project).

[6] See Kitcher 1992, p.101.

Who are the proponents of complaint (C)? Clearly one would expect Kitcher's "radical naturalists," particularly sociologists of science, to support this complaint and this is indeed what we find:

> One prominent form of contemporary naturalism, popular among many historians and sociologists of science, appeals to (C) to support the conclusion that normative epistemology is an exercise in empty moralizing. (Kitcher 1992, p.80)

Kitcher's villains here are the usual suspects of contemporary sociological (or constructivist) relativism, e.g. Barnes, Bloor, Shapin, Schaffer, Collins, Latour and Feyerabend. There is nothing surprising about Kitcher's objection to this use of complaint (C) nor should it surprise us that Kitcher should name and shame such outspoken critics of normative philosophy of science.

However, what should surprise us is that he should choose to include among this crowd the work of Laudan.[7] This inclusion is puzzling for two reasons. Firstly, it is the *only* mention of Laudan's work in a paper that is chiefly concerned with understanding the development of contemporary naturalism. This is puzzling because Laudan was one of the most prominent and able proponents of naturalism in the philosophy of science from the late 1970s to the early 1990s.[8] Secondly, and more importantly, it seems strange to group Laudan's views on the transitory nature of scientific goals with those of contemporary relativists. In contrast to such relativism, Laudan (1984) is at pains to point out that the absence of a single conception of cognitive virtue need not lead down the road to relativism but is instead fully compatible with the development of a normative, progressive and above all rational account of science. Indeed, Laudan (1981) presents a sustained attack on what he sees as the shallow and 'pseudo-scientific' approach of Bloor (1976) and other adherents of the Strong Programme in the sociology of science.

Of course, Kitcher is fully aware of the fact that Laudan's position is neither relativist nor anti-normative. However, to acknowledge this fact is only to make it even more puzzling that Kitcher did not emphasize the quite different roles complaint (C) can play vis-à-vis traditional naturalism. The explanation for this rather odd classification of Laudan's normative naturalism is that Kitcher takes the specification of a universal goal for inquiry to be constitutive of traditional naturalism. Because both sociological

---

[7] See Kitcher 1992, p.106.

[8] See chapter 1.

relativists and Laudan insist that no such goal exists, Kitcher sees them both as denying the possibility of traditional naturalism. On this view, it does not matter that Laudan's axiological relativism is part of an account of progress and rationality in science; a commitment to the specification of a universal goal for inquiry is a necessary condition for being a traditional naturalist. In contrast, this is *not* a necessary condition for Rosenberg's naturalized philosopher of science. This is because Rosenberg, pace Kitcher, does not take the specification of a universal goal for inquiry as being constitutive of naturalized philosophy. Rather, he claims that a shared commitment to progressivity can be cashed out in several different ways; some of which specify truth as the goal of science whereas others do not.

How are we to decide which of these competing accounts is the most accurate specification of naturalism in the philosophy of science? From what has been said so far it may appear that Rosenberg's account is more accurate in the sense that it avoids lumping Laudan's normative naturalism together with the anti-normative naturalism of sociological relativism. However, I want to suggest that *neither* Kitcher's account nor Rosenberg's is a more accurate description of naturalized philosophy of science and its relationship to competing positions. Indeed, it seems to me that questions of 'accuracy' are misplaced here and instead we should ask which account is more 'appropriate' for a particular explanatory or investigative purpose. If like Kitcher one is interested in naturalism as a way of developing a naturalistic epistemology that retains the traditional goals of epistemology then it is *appropriate* to view positions like Laudan's as part of a cluster of radical positions that challenge this melioristic project. However, if like Rosenberg one is interested in naturalism as a way of developing a realist account of progress in science then it is *appropriate* to see Laudan as providing an *internal* critique of the view that progressivity requires us to specify a single, universal goal for cognitive inquiry.

So, given that we are interested in the development of naturalistic realism it would seem that Rosenberg's account is more appropriate for our explanatory purposes. Thus, in the rest of this chapter I will show how contemporary arguments for naturalistic realism must be understood in terms of how they react to three separate positions that all (in one way or another) deny the possibility of defending realism from a naturalistic perspective: Laudan's naturalistic antirealism, van Fraassen's constructive empiricism, and sociological relativism.

### 3. Laudan's naturalistic antirealism

In *Progress and its Problems*, Laudan put forward an antirealist interpretation of science according to which scientific progress is judged in terms of problem-solving effectiveness rather than the standard model of moving closer and closer to an accurate representation of the world. However, it is notable that beyond the claim that truth is a utopian (i.e. unattainable) goal, Laudan (1977) offers no sustained argument against a realistic interpretation of science.[9] This changed with the publication of Laudan's seminal paper 'A Confutation of Convergent Realism' in which he attempts to refute a new kind of argument for scientific realism:

> A growing number of philosophers (including Boyd, W. Newton-Smith, A. Shimony, Putnam, and I. Niiniluoto) have argued that the theses of epistemic realism are open to empirical test. The suggestion that epistemological doctrines have much the same empirical status as the sciences is a welcome one; for, whether it stands up to detailed scrutiny or not, this suggestion marks a significant facing-up by the philosophical community to one of the most neglected (and most notorious) problems of philosophy: the status of epistemological claims. (Laudan 1996a, p.107)

The position in question here models its argument for realism on a standard form of abductive reasoning found in science. The data to be explained is the apparent success of science and the competing explanations of this data are rival philosophical theories of science. The naturalistic realist then claims that realism provides the best explanation of why science should be so strikingly successful. This is the so-called *explanationist defence* of scientific realism.

Laudan (1981/1996, 1984) begins by suggesting that the explanationist defence of scientific realism is based on the attempt to establish variants of the following claims:

(CER-1) Scientific theories (at least in the 'mature sciences') are typically approximately true, and more recent theories are closer to the truth than older theories in the same domain.

(CER-2) The observational and theoretical terms within the theories of a mature science genuinely refer.

---

[9] Rather, his early rejection of realism stems from a belief that the truth or falsity of scientific theories is largely *irrelevant* to an assessment of their cognitive value. See Laudan (1977), p.23 and p.120.

(CER-3) Successive theories in any mature science will be such that they preserve the theoretical relations and the apparent referents of earlier theories; that is, earlier theories will be limiting cases of later theories.

(CER-4) Acceptable new theories do and should explain why their predecessors were successful in so far as they were successful.

(CER-5) Theses CER-1 to CER-4 entail that ('mature') scientific theories should be successful; indeed, these theses constitute the best, if not the only, explanation for the success of science.

Theses CER-1 to CER-5 constitute Laudan's understanding of the argument advocated by Boyd, Putnam and Newton-Smith whereby the explanationist/miracle manoeuvre (CER-5) is invoked to defend a particular version of scientific realism (CER-1 to CER-4). Laudan calls the resulting position *convergent epistemological realism.*

It should come as no surprise that although Laudan applauds this step toward naturalism he rejects the explanationist defence of scientific realism. Thus, he argues that the history of science fails to support the claim that realism is the best explanation of the success of science. Further, he suggests that the large number of successful theories that turned out to be false provides grounds for a pessimistic induction according to which we have reason to think that current successful theories are also false. Although Laudan (1984) subsequently attempted to incorporate these arguments into his 'reticulated model' of scientific rationality, it is the earlier critique of convergent epistemological realism that has proved to be his lasting legacy. For this reason I have chosen to focus exclusively on this (negative) aspect of Laudan's naturalistic antirealism. This is not to say that his positive proposals concerning scientific rationality are obviously flawed or irrelevant, it is simply a reflection of our explanatory interests. If one is interested in understanding contemporary arguments for naturalistic realism then it is the pessimistic induction rather than the reticulated model that needs to be considered.

### 3.1 The argument against synchronic CER

Laudan suggests that proponents of convergent epistemological realism rely on two abductive arguments for their claim (CER-5) that theses CER-1 to CER-4 constitute the best explanation of science. The first of these arguments attempts to show that approximate truth (CER-1) and reference (CER-2) are at least past of such an explanation:

(ATR-1) If scientific theories are approximately true, then they typically will be empirically successful.

(ATR-2) If the central terms in scientific theories genuinely refer, then those theories generally will be empirically successful.

(ATR-3) Scientific theories are empirically successful.

(ATR-4) (Probably) theories are approximately true and their terms genuinely refer.

Laudan's aim is to show that ATR-1 and ATR-2 are false, thereby blocking the realist's desired conclusion that reference and approximate truth feature in the best explanation of the success of science. Let us start by looking at Laudan's argument against the realist's attempt to hook up reference and success.

Laudan attributes the referential side of the explanationist defence of scientific realism to Putnam (1978). For Laudan, the only way Putnam can establish the connections between reference and success that the explanationist defence requires is to support the following chain of inference:

(RS-1) The theories in the advanced or mature sciences are successful.

(RS-2) A theory whose central terms genuinely refer will be a successful theory.

(RS-3) If a theory is successful, we can reasonably infer that its central terms refer.

(RS-4) All the central terms in theories in the mature sciences do refer.

It should be noted that there are two arguments working together here; one that shows why reference explains or leads to success (i.e. RS-4 with RS-2 explains RS-1), and another that shows why success can be taken as grounds for reference (i.e. RS-1 with RS-3 allows us to infer RS-4). Laudan's claim is that neither of these arguments (or for that matter their constituent claims) is acceptable given the historical record of science.

In order to show this Laudan presents a number of examples from the history of science that would seem to undermine the connections between reference and success postulated by RS-2 and RS-3. Thus, against the claim that genuinely referential theories are invariably successful (RS-2) Laudan suggests the following counterexamples:

- Chemical atomic theory in the 1700s.
- Proutian theory in the 1800s.
- Wegnerian theory from the 1920s to the early 1960s.

Laudan's point is that although notably unsuccessful in the periods indicated all of these theories postulated entities that we would now take to genuinely exist. His explanation for this seemingly counterintuitive claim being that although these theories referred to genuinely existing entities they failed to identify the properties of these entities or their modes of interacting. Once this is recognized we are invited to conclude that there is no reason to think that securing reference ensures empirical success, i.e. RS-2 is false.[10]

Laudan's argument against RS-3 also uses historical evidence to undermine the claim that the empirical success of a scientific theory can be taken as grounds for the belief that its central terms genuinely refer.[11] Laudan's strategy is the opposite of that employed against RS-2 for now he presents a series of counterexamples where we have empirically successful theories whose central terms *do not* refer. Such counterexamples centre on the role of the (non-referring) ether in scientific theories of the 1830s and 1840s:

- The caloric ether in chemistry and heat theory.
- The optical ether in the theory of light.
- Gravitational and physiological ethers.

Again, Laudan's point is that although non-referential the concept of the ether played a central role in many undeniably successful scientific theories of the nineteenth century. Laudan's worry is that such counterexamples seems to force the realist into making one of two claims; either they can claim that ethereal theories were not successful because they employed a non-referential term or that 'ether' *is* a genuinely referring term. The first claim makes the connection between success and reference trivial (or ad hoc)

[10] Laudan also shows that any attempt to weaken RS-2 will also be subject to refutation by the historical record (see Laudan 1996a, p.113).
[11] The seeds of this argument can be found in Laudan's earlier work (see Laudan 1977, p.82).

whereas the second requires us to believe in the existence of entities rejected by modern science. Because neither of these claims is acceptable Laudan suggests that RS-3 is false.[12]

Having dismissed the referential side of the explanationist defence Laudan turns to the attempt to defend realism via the connection between success and approximate truth (ATR-1). Laudan suggests that the attempt to connect approximate truth and success is best characterized as a commitment to the following claims:

(TS-1) If a theory is approximately true, then it will be explanatory successful.

(TS-2) If a theory is explanatorily successful, then it is probably approximately true.

Laudan's problem with TS-1 is that although there is an intuitive connection between truth and success, it is far harder to establish a genuine entailment between approximate truth and success.[13] He says this primarily because of the lack of an adequate definition of what it means for something to be approximately true. Laudan challenges the realist to supply such a definition, rejecting the claim that the notion of approximate truth can legitimately be invoked without it.[14]

The argument against TS-2 sees us return to familiar territory because the strategy employed against it is exactly the same one Laudan uses against RS-3, the claim that we can infer reference on the basis of empirical success. Laudan is able to do this because of the connection that would seem to exist between reference and approximate truth:

> I take it that a realist would never want to say that a theory was approximately true if its central terms failed to refer...In short, a necessary condition, especially for the scientific realist, for a theory being close to the truth is that its central explanatory terms genuinely refer. (Laudan 1996a, p.121)

---

[12] Laudan argues that any attempt to weaken RS-3 by restricting it to certain parts of the theory will remove the rationale for the realist's claims about convergence, retention, and correspondence in inter-theory relations (see Laudan 1996a, p.117).

[13] Laudan correctly points out that most realists would like to be able to substitute 'true' for 'approximately true' in this statement but that they are reluctant to say this partly because of their belief that science is likely to revise our current evaluations of truth (see Laudan 1996a, p.118).

[14] As Laudan says, it is not the concept of approximate truth itself we are interested in but rather its role in the inference from approximate truth to success.

This is a very clever move by Laudan because all of the counterexamples against RS-3 (a claim concerning the connection between reference and success) are now just as relevant for assessing TS-2 (a claim concerning the connection between approximate truth and success). However, instead of focussing on the previously discussed nineteenth century ethereal theories, Laudan now produces a long list of successful yet non-referential scientific theories:

- The crystalline spheres of ancient and medieval astronomy.
- The humoral theory of medicine.
- The effluvial theory of static electricity.
- 'Catastrophist' geology, with its commitment to a universal (Noachian) deluge.
- The phlogiston theory of chemistry.
- The caloric theory of heat.
- The vibration theory of heat.
- The vital force theories of physiology.
- The electromagnetic ether.
- The theory of circular inertia.
- Theories of spontaneous generation.

Just how successful some of these 'theories' were is certainly open to question but this and other issues concerning Laudan's choice of counterexamples will be discussed in the next chapter. For the moment let us concentrate on Laudan's response to a typical realist way of avoiding the problems alluded to in this section.

Faced with the counterexamples presented by Laudan the realist may choose to claim that their account is only applicable to the 'mature sciences'. As Laudan notes:

> This distinction between mature and immature sciences proves convenient to the realist, since he can use it to dismiss any prima-facie counter-example to the empirical claims of CER on the grounds that the example is drawn from a so-called immature science. (Laudan 1996a, p.122)

For Laudan, there are two problems with this approach. The first and most obvious problem is that it makes CER completely ad hoc because an "immature" science will turn out to be anything that doesn't fit in with the CER analysis. The second more interesting problem is that the claim to be dealing with only the mature sciences seems to put unnecessary restrictions on the realist's avowed explanandum, namely the success of science. If Laudan is right then the realist will end up explaining the success of a very small amount of what we would generally call 'successful science'. In other words, the 'mature sciences' gambit would seem to undermine the whole rationale behind the so-called 'miracle argument' for now there will be large parts of science whose success is unexplained and therefore remains 'miraculous'.[15]

### 3.2 The argument against diachronic CER

So far we have only discussed Laudan's arguments against the synchronic form of realism (CER-1, CER-2 and CER-5) whose basic argumentative structure (ATR-1 to ATR-4) concerns the related notions of reference (RS-1 to RS-4) and approximate truth (T-1 and T-2).[16] If the realist wishes to defend a strong diachronic account of realism (i.e. to include CER-3 and CER-4) Laudan suggests that he must also commit himself to the following claims:

(CC-1) If earlier theories in a scientific domain are successful and thereby, according to realist principles, approximately true, then scientists should only accept later theories that retain appropriate portions of earlier theories.

(CC-2) As a matter of fact, scientists do adopt the strategy of (CC-1) and manage to produce new, more successful theories in the process.

(CC-3) The 'fact' that scientists succeed at retaining appropriate parts of earlier theories in more successful successors shows that the earlier theories did genuinely refer and that they are approximately true. And thus, the strategy propounded in (CC-1) is sound.

The basic idea expressed in these claims is that scientific progress is cumulative because later theories preserve the successful parts of earlier theories, and that (again) the best explanation of this fact is the realist's account of reference and approximate truth.

---

[15] Laudan also points out that the 'mature sciences' gambit will make many recent examples of successful science immature (see Laudan 1996a, p.125).

[16] If this seems confusing it should be pointed out that Laudan's (1981/1996) way of presenting his argument hardly lends itself to simple explication.

Against the strong diachronic form of CER Laudan attempts to establish two important features of science that seem to conflict with claims CC-1 and CC-2:

1. As a matter of fact scientists *do not* adopt the strategy of accepting later theories only when they preserve appropriate portions of earlier theories, i.e. CC-2 is false.
2. As a matter of fact later theories *are not* typically limiting cases of earlier theories, i.e. CC-1 is not sound methodological advice.

Laudan's grounds for making these claims again rely on his selection of certain counterexamples. Thus, in order to provide grounds for statement 1, Laudan gives the following examples:

- The transformation from the corpuscular to the wave theory of light.
- The transformation from catastrophist to uniformitarian geology.
- The transformation from Lamarckian to Darwinian evolutionary theory.

Laudan's point is that if the scientists in the above episodes had followed the retentionist strategy advocated by the realist then we would expect these transformations to be criticized due to their non-cumulative nature. The 'fact' that there were no such criticisms shows, for Laudan at least, that scientists did not adopt the retentionist strategy in these cases.

So Laudan thinks he has shown that, as a matter of fact, the retentionist strategy has not been employed in several important cases of theory change in the history of science. However, as Laudan notes, it could still be the case that theories themselves illustrate the kind of relationship required by the realist's account of cumulation and convergence. Against this response Laudan points to the following cases:

- Copernican astronomy did not retain all the key mechanisms of Ptolemaic astronomy.
- Newtonian physics did not retain all the theoretical laws of Cartesian mechanics, astronomy, and optics.
- Franklin's electrical theory did not contain its predecessor as a limiting case.

- Relativistic physics did not retain the ether, nor the mechanisms associated with it.
- Statistical mechanics does not incorporate all the mechanisms of thermodynamics.
- Modern genetics does not have Darwinian pangenesis as a limiting case.
- The wave theory of light did not appropriate the mechanisms of corpuscular optics.
- Modern embryology incorporates few of the mechanisms prominent in classical embryological theory.

It is difficult to see why the realist could not claim that in all of these cases the later theories have retained the successful parts of the earlier theories and rejected the rest. For example, it seems obvious to say that the reason why modern genetics did not retain Darwinian pangenesis was for the simple fact that it was wrong.

However, this response misses the essential point of Laudan's argument for his target is precisely the realist's conservative attachment to the ontological framework of modern science:

> In spite of his commitment to the growth of knowledge, the realist would unwittingly freeze science in its present state by forcing all future theories to accommodate the ontology of contemporary (mature) science and by foreclosing the possibility that some future generation may come to the conclusion that some (or even most) of the central terms in our best theories are no more referential than was 'natural place', 'phlogiston', 'ether', or 'caloric'. (Laudan 1996a, p.131)

So, if Laudan is right, the realist is left with a rather unpalatable dilemma: either to stick to the "limiting case" model of scientific progress in the face of overwhelming historical evidence, or accept the fact that there have been deep ontological shifts in the history of theory change thereby weakening our belief that the central terms of our current theories genuinely refer. As we shall see in section 5, most realists have attempted to steer a course somewhere between these two horns by advocating a cumulative model of scientific progress that allows for significant changes in ontology.

The effect of Laudan's empirical argument against strong diachronic CER is to undercut the abductive inference in CC-3 by showing that there is little reason to accept the realist's alleged explanandum, namely that later theories preserve earlier theories as limiting cases.[17] However, Laudan concedes that it may be possible to defend a weaker version of diachronic CER that it is based on CER-4, namely the claim that later theories should be able to explain why their predecessors were successful without necessarily preserving their content. Laudan's reasons for rejecting this weakened form of CER are threefold. Firstly, it is gratuitous because the ability of a theory to explain the success of its predecessor is neither a necessary nor a sufficient condition for saying that it is better. Secondly, the notion of "explaining" being used here is too ambiguous to provide support for realism. Laudan concludes that the weakened form of CER is no better off than the other forms he considers.

### 3.3 The methodological critique of CER

Laudan's first two arguments challenge the realist to explain the seemingly endless number of cases that fail to exhibit the kind of characteristics required if the explanationist defence is to go through. In this sense Laudan is attacking the *content* of the realist's argument rather than its basic inferential structure. However, like van Fraassen (1980) and Fine (1986a), Laudan (1996a) also argues that the very *form* that the explanationist defence of scientific realism takes begs the question against alternative explanations of the success of science. For in order to show that the best explanation of the success of science is the approximate truth or genuine referential capacity of scientific theories the realist must assume that the success of a theory is good reason for believing it to be true. However, as Laudan points out, this last assumption is exactly what antirealists want to deny:

> The non-realist refuses to admit that a *scientific* theory can be warrantedly judged to be true simply because it has some true consequences. Such non-realists are not likely to be impressed by the claim that a philosophical theory such as realism can be warranted as true because it arguably has some true consequences. If non-realists are chary about first-order abductions to avowedly true conclusions, they are not likely to be impressed by second-order abductions. (Laudan 1996a, p.134)

Here we can see exactly why the explanationist defence is so controversial for it seems to use exactly the same form of inference at the philosophical level that its opponents dispute at the scientific level. As Laudan notes, given this fact it is hard to see how this line of reasoning could convince anyone other than a realist.

---

[17] Laudan also attempts to show that the realist's account of cumulation is, in principle, unattainable. He attributes this argument to David Miller (see Laudan 1996a, p.131-2).

However, there is the possibility that the explanationist defence is not in fact intended to convince the committed antirealist but is rather an attempt to show that realism is at least as well confirmed as any other putative explanation of science. The motivation behind this suggestion being that it licences the use of inferential principles that only realists would accept. As Laudan sees it the problem with this weakening of the explanationist defence is that it although it may avoid the charge of begging the question it relies on a rather implausible account of confirmation:

> Since realism was devised to explain the success of science, it remains purely *ad hoc* with respect to that success. If realism made some novel predictions or has been subjected to carefully controlled tests, one does not learn about it from the literature of contemporary realism. (Laudan 1996a, p.135)

Laudan's point here is that in advocating the weak version of the explanationist defence the realist seems to be prepared to accept a form of *ad hoc* explanation at the philosophical level that they are quite opposed to at the scientific level.

## 4. Van Fraassen's constructive empiricism

Bas van Fraassen is perhaps the most famous defender of empiricism in the philosophy of science. Like Laudan, van Fraassen argues that although our theories may be literally true or false we can never know that they are true or false (i.e. he is a semantic realist but not an epistemological realist). For example, although he accepts the Kuhnian critique of logical positivism (Kuhn 1970), he insists that we must not throw the empiricist baby out with the positivist bathwater. As Giere (1988) says:

> Van Fraassen rejected the logical empiricist view of theories as interpreted formal systems. He rejected any attempt to construct an inductive logic. And he rejected the logical analysis of scientific explanation. Nevertheless, van Fraassen's positive account of science upholds the deflationary anti-realism that inspired logical empiricism. (Giere 1988, p.47)

So, for van Fraassen, the chief flaw of logical positivism was an over preoccupation with linguistic issues, *not* its empiricism. Indeed, van Fraassen suggests that the empiricism of Mach, Duhem and Poincaré can be resurrected if only we can correct the flaws in the logical positivist account. Thus, van Fraassen's aim is to combine an empiricist argument for antirealism with a non-linguistic approach to issues like theory structure, modality and the observable-unobservable distinction. His claim that we can do this is the foundation of *constructive empiricism*, the view that we should only accept scientific theories as empirically adequate descriptions of observable phenomena.

Van Fraassen makes his case for constructive empiricism by answering three questions:

1. How do theories relate to the world?
2. What is a scientific theory?
3. What is the best explanation of science?

Although there is a very close connection between these three questions, for ease of exposition it is convenient to treat van Fraassen's answers to them separately. Roughly speaking, his answer to the first question is what constitutes the *empiricist* aspect of constructive empiricism whereas his answer to the second question is what constitutes its *constructive* aspect. Finally, his answer to the third question is his case for why we should prefer constructive empiricism to its rivals. Let us begin then with the case for empiricism.

### 4.1 How do scientific theories relate to the world?

In order to understand van Fraassen's empiricism one must start with his characterization of scientific realism. Rejecting a naïve formulation of scientific realism on the grounds that it commits realists to the claim that all *current* scientific theories are true, van Fraassen offers the following minimal definition of scientific realism:

> Science aims to us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true (van Fraassen 1980, p.8)

Van Fraassen argues that this definition is one that any scientific realist worthy of the name would agree to. However, the real purpose of the definition is to allow van Fraassen to draw a distinction between two varieties of antirealism:

> The idea of a literally true account has two aspects: the language is to be literally construed; and so construed, the account is true. This divides anti-realists into two sorts. The first sort holds that science is or aims to be true, properly (but not literally) construed. The second holds that the language of science should be literally construed, but its theories need not be true to be good. (van Fraassen 1980, p.10)

As van Fraassen notes, the first sort of antirealism is familiar from positivist and instrumentalist accounts of science where, it is argued, claims concerning theoretical entities are *really* about sense-data or observable regularities and when properly construed in this way are 'true'. In contrast, the second sort of antirealism, the one that van Fraassen supports, interprets the language of science literally. For van Fraassen, if a physical theory claims that electrons exist then that is exactly how we should understand it, "the apparent statements of science really are statements, *capable of* being true or false" (van Fraassen 1980, p.10). This is just the first of several ways in which van Fraassen's antirealism breaks with its empiricist predecessors.

Of course, to say that scientific theories must be literally construed is not to say that we must accept what they say as true. As van Fraassen says, "this insistence relates not at all to our epistemic attitudes toward theories, nor to the aim we pursue in constructing theories, but only to the correct understanding of *what a theory says*" (van Fraassen 1980, p.11). In fact, van Fraassen argues that we *never* have reason to believe that the entities postulated by scientific theories actually exist. This is the claim at the heart of constructive empiricism, according to which:

> Science aims to give us theories which are empirically adequate and acceptance of a theory involves as belief only that it is empirically adequate. (van Fraassen 1980, p.12)

To fully appreciate the notion of *empirical adequacy* one needs to understand van Fraassen's account of theory structure. However, for the moment we can focus on the fact that, for van Fraassen, to accept a theory is simply to believe that it 'saves the phenomena.' Thus, in contrast to the realist, van Fraassen's constructive empiricist is prepared to accept a theory purely on the grounds that what it says about the observable aspects of the world is true.[18] As a corollary, what the theory says about unobservable aspects of the world turns out to be irrelevant to the acceptance of the theory in question.

Clearly, if one is going to suggest that we should only accept what a theory says about observable phenomena we must have some way of distinguishing between observable and unobservable features of the world. Of course, the difficulty in carrying

[18] As it stands this claim is strictly speaking incorrect. For van Fraassen, we accept a theory for its empirical adequacy *and* certain pragmatic virtues. I have chosen to ignore this aspect of van Fraassen's account on the grounds that it does not bear heavily on the issues of this chapter, i.e. the realism issue.

out this task was amply demonstrated by the failure of the logical positivists to construct a pure observation language that would be linked to the theoretical postulates of science by means of correspondence rules. However, van Fraassen does not think that the failure of this project should be taken to show that there is no way of drawing a meaningful distinction between the theoretical and the observable. In fact, as van Fraassen suggests, there are actually two issues at stake here: the division of *language* into a theoretical and non-theoretical part, and the division of *objects and events* into observable and unobservable. As far as the first issue is concerned van Fraassen agrees with Sellars and Feyerabend that "all our language is theory-infected" (van Fraassen 1980, p.14), there is no way of drawing a distinction between theoretical and non-theoretical aspects of our language. However, as far as the second issue is concerned, van Fraassen does not accept that there is no meaningful way of distinguishing between observable and unobservable objects and events.

In order to understand van Fraassen's argument for this claim one must again appreciate the difference between what a theory says and how much of that theory we choose to believe:

> The fact that we let our language be guided by a given picture, at some point, does not show how much we believe about that picture...no immediate conclusions can be drawn from the theory-ladenness of our language. (van Fraassen 1980, p.15)

In particular, van Fraassen urges us not to take the lack of a distinction between the theoretical and non-theoretical as grounds for rejecting the possibility of drawing a distinction between the observable and the unobservable. A separate argument is required to show the impossibility of drawing the latter distinction. Van Fraasen considers three such arguments, all of them due to Maxwell (1962).

The first of these arguments concerns the continuum of cases that exists between direct observation and inference. Here Maxwell argues that because there is a continuum from the observable to the unobservable (i.e. looking through a window, looking through glasses, looking through binoculars, looking through a low-powered microscope, looking through a high-powered microscope, etc.) there is no way of drawing a non-arbitrary distinction between the theoretical and the observable. The second argument suggests that anything can count as observable given different circumstances, e.g. a possible world in which we have microscope eyes would contain very different 'observable' objects. The third argument picks up on this arcane suggestion by questioning the usefulness of the observable-unobservable distinction even if it could be drawn. Here it is suggested that being "an accident and a function of our physiological make-up" (Maxwell quoted in van Fraassen 1980, p.18) any formulation of the observable-unobservable distinction can have no ontological significance because it has nothing to do with the existence of the objects and events in question.

Van Fraassen's responses to these arguments are intriguing and at times unexpected, but they again serve to illustrate just how far he is prepared to go to defend an empiricist philosophy of science. Thus, in response to the first argument, van Fraassen admits that there is a continuum from the observable to the unobservable but denies that this is of any particular significance. Instead, he suggests that it merely demonstrates that 'observable,' like many other predicates, is *vague*. In van Fraassen's view, vague predicates are perfectly usable "provided there are clear cases and clear counter-cases" (van Fraassen 1980, p.16), which he argues are both readily available for a term like 'observable'.

In response to the suggestion that anything can count as 'observable' because, among other things, we might have electron microscope eyes, van Fraassen says:

> This strikes me as a trick, a change in the subject of discussion. I have a mortar and pestle made of copper and weighing about a kilo. Should I call it breakable because a giant could break it? Should I call the Empire State Building portable? Is there no distinction between a portable and a console record player? (van Fraassen 1980, p.17)

Van Fraassen's reductio of Maxwell's argument is highly convincing. After all, if we cannot define 'observable' because we might have electron microscope eyes, what is to stop us from objecting to any putative definition of a predicate using 'if only' counterfactual conditions? Shouldn't we allow certain limitations to fix the meaning of particular kinds of predicates? As van Fraassen says:

> The human organism is, from the point of view of physics, a certain kind of measuring apparatus. As such it has certain inherent limitations – which will be described in detail in the final physics and biology. It is these limitations to which the 'able' in 'observable' refers – our limitations, *qua* human beings. (van Fraassen 1980, p.17)

So, the meaning of predicates like 'observable' and 'breakable' is dependent on certain facts about the kinds of speakers who use them. On this view, there is no problem with justifying talk about observable objects because 'observable' simply means 'observable-to-us'.

This claim that the notion of observability is relative to the cognitive abilities of human observers is also at the heart of van Fraassen's response to Maxwell's third argument. Here van Fraassen suggests that we can admit that observability is an entirely anthropocentric construction without undermining the case for empiricism:

> It is, on the face of it, not irrational to commit oneself only to a search for theories that are empirically adequate, ones whose models fit the observable phenomena, while recognizing that what counts as an observable phenomenon is a function of what the epistemic community is (that *observable* is *observable-to-us*). (van Fraasen 1980, p.19)

Again, van Fraassen's point here is that one should not confuse ontological issues concerning existence with the appropriate epistemic attitudes to scientific theories. As we have seen, van Fraassen is perfectly prepared to allow that unobservable theoretical entities may exist; he just doesn't think we can ever have good reason to believe that they do exist. If van Fraassen is right the most pressing issues in science concern what we can *know* to exist, not what might exist independently of whether or not we know that it does.

### 4.2 What are scientific theories?

Having defended the coherence of an empiricist account of science van Fraassen moves on to the question of what scientific theories actually are and how this bears on the possibility of outlining an empiricist alternative to the logical positivists' account of theory structure. For van Fraassen this requires us to answer one particular question: what is the empirical content of a theory?

Van Fraassen begins by distinguishing two possible approaches to this question. The syntactic approach of the logical positivists suggests that theories are linguistic entities that are defined in terms of a finite set of axioms or laws. The empirical content of a theory is then defined in terms of the empirical consequences of these axioms or laws that are linked to an observation language by correspondence rules. In contrast, the semantic approach to theory structure suggests that it is better to think of theories as a set or family of *models* that satisfy a certain set of axioms or laws. For van Fraassen, the significance of this approach is that it allows us to shift attention away from the linguistic focus of the syntactic approach:

> The syntactic picture of a theory identifies it with a body of theorems, stated in one particular language chosen for the expression of that theory. This should be contrasted with the alternative of presenting a theory in the first instance by identifying a class of structures as its models. In this second, semantic, approach, the language used to express the theory is neither basic nor unique; the same class of structures could well be described in radically different ways, each with its own limitations. The models occupy centre stage. (van Fraassen 1980, p.44)

Van Fraassen finds support for the semantic approach in the fact that scientists often speak of 'models,' although he does note that their usage of this term differs slightly from the meaning this term is given in logic and meta-mathematics. If the semantic approach is favoured over the syntactic approach the next important question is this: if non-linguistic models rather than language take centre in science what is the empirical content of a model?

To appreciate van Fraassen's answer to this question one must return to the origins of the concept of a model in geometry and logic. Here van Fraassen shows that a crucial feature of geometrical models is that they can be *embedded* in each other:

> We say that one structure can be embedded in another, if the first is isomorphic to a part (substructure) of the second. Isomorphism is of course total identity of structure and is a limiting case of embeddability: if two structures are isomorphic then each can be embedded in the other. (van Fraassen 1980, p.43)

The key move in van Fraassen's empiricist account of theory structure is the claim that the isomorphic relationship that exists between geometric models is also the best way of understanding the relationship that exists between scientific models and the world. Thus, in van Fraassen's view, a theory is empirically adequate if has one model that is isomorphic to the observable features of the world.[19] This is the technical meaning of van Fraassen's claim that we should only *accept* theories as being empirically adequate descriptions of observable phenomena rather than the *belief* that they are true representations of the world.

Why does van Fraassen think that empiricism is better couched in terms of the semantic approach rather than the syntactic approach? His answer to this question

[19] He introduces this point by means of an example that focuses on Newton's views on absolute space. See van Fraassen 1980, pp.44-47.

concerns the way in which these two approaches explicate the related concepts of *empirical adequacy* and *empirical equivalence*. As we have seen, according to van Fraassen's semantic approach, to claim empirical adequacy for a theory is to suggest that it has a model that is isomorphic to the observable phenomena (or that it 'saves the phenomena'). Similarly, to say that two theories $T_1$ and $T_2$ are empirically equivalent is to say that the observable consequences of a model in $T_1$ are isomorphic to the observable consequences of a model in $T_2$. Van Fraassen's claim is that we should prefer these semantic accounts of empirical adequacy and equivalence because attempts to explicate them syntactically have "conspicuously failed" (van Fraassen 1980, p.53).

### 4.3 What is the best explanation of science?

Even if van Fraassen can defend his semantic version of empiricism from the technical difficulties that plagued the logical positivists' syntactic account of theories he still needs to show why it constitutes a better account of science than realist alternatives. In particular, he needs to respond to the claim that realism, rather than constructive empiricism, is the *best explanation* of science. Here van Fraassen begins by suggesting that this sort of argument for realism is based on an appeal to the way we reason in 'ordinary' situations:

> It is argued that if we follow this rule in all 'ordinary' cases; and that if we follow it consistently everywhere, we shall be led to scientific realism...And surely there are many telling 'ordinary' cases: I hear scratching in the wall, the patter of little feet at midnight, my cheese disappears – and I infer that a mouse has come to live with me. Not merely that these apparent signs of mousely presence will continue, not merely that all the observable phenomena will be as if there is a mouse; but that there really is a mouse. (van Fraassen 1980, pp.19-20)

The realist wants to say that just as the actual existence of a mouse is the best explanation of scratching in the wall and the patter of little feet, so too is the existence of unobservable entities the best explanation of science. In response to this argument, van Fraassen offers two objections.

Firstly, he argues that because it is an *empirical* claim that we use inference to the best explanation to infer the presence of the mouse it is perfectly legitimate to present rival hypotheses. So, van Fraassen suggests that our ability to infer the presence of mice is equally well explained by the rival hypothesis that "we are always willing to believe that the theory which best explains the evidence, is empirically adequate (that all the observable phenomena are as the theory says they are)" (van Fraassen 1980, p.20).

The problem for the realist is that because mice *are* observable there is no way of deciding between the realist hypothesis that we follow the rule of inference to the best explanation in ordinary situations and the empiricist alternative that claims that we don't. The presence of the mouse provides no telling evidence in favour of one hypothesis rather than the other. His second objection is that even if we accept inference to the best explanation it cannot be used to rule out alternative antirealist explanations without a further premise:

> The realist asks us to choose between different hypotheses that explain the regularities in certain ways; but his opponent always wishes to choose among hypotheses of the form 'Theory $T_i$ is empirically adequate'. So the realist will need his special extra premiss that every universal regularity in nature needs an explanation, before the rule will make realists of us all. And that is just the premiss that distinguished the realist from his opponent. (van Fraassen 1980, p.21)

So again the problem for the realist is that he cannot show that realism is the best explanation of science without *begging the question* against alternative antirealist hypotheses, i.e. by assuming that all universal regularities in nature are in need of explanation. Further, van Fraassen argues that various attempts to circumvent this problem by providing independent warrant for the "special extra premise" (e.g. Smart's 'no cosmic coincidences' argument, Reichenbach's principle of the common cause, and Putnam's 'no miracles' argument) all repeat the same basic question begging manoeuvre.

So van Fraassen concludes that the realist will not be able to show that realism is the best explanation of the success of science without begging the question against alternative antirealist explanations. In fact, he argues that the realist's confidence in her own explanation as obviously superior to such alternatives is misplaced because there is a much better answer to the question of why we have successful theories in science. If science is seen as a biological phenomenon then it is possible to explain the success of science without appealing to unobservable entities or processes:

> [The] success of science is no miracle. It is not even surprising to the scientific (Darwinist) mind. For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive – the ones which in fact latched on to regularities in nature. (van Fraassen 1980, p.40)

Here then we have an *evolutionary* justification for rejecting the realist's demand for 'deep' explanations of the success of science. On this view, just as we explain the survival of certain species in terms of their ability to adapt to their environment so too can we explain the survival of certain scientific theories in terms of their ability to latch on to (observable) regularities in nature. No further explanation is required. Thus, for van Fraassen, we should reject realism on the grounds that it attempts to provide an explanation that scientists themselves would deem to be unnecessary.

## 5. SSK and social constructivism

Because the pessimistic induction and constructive empiricism are the products of individual philosophers it is relatively easy to summarise their respective claims and arguments. Unfortunately, this is certainly not the case with the third major opponent of naturalistic realism, the sociology of science. The problem here is that there are almost as many sociologists of science as there are sociologies of science. However, it is generally agreed that contemporary sociology of science is best characterised in terms of its commitment to the idea of 'social construction' (Fine 1996; Golinski 1998; Kukla 2000; Hacking 1999). As Fine (1996) says:

> Until recently social-constructivist ideas had not been central to investigations in the sociology of science, which was dominated instead by the institutional approach of the Merton school. This situation has changed over the past decade or so with the rise of a movement, referred to by its practitioners as the "sociology of scientific knowledge," that incorporates the idea of social construction in an especially striking way. (Fine 1996, p.231)

Although some commentators have a tendency to use the term "social constructivism" to refer to pretty much every major development in the sociology of science over the last thirty years, it is important to realise that the claim that science is socially constructed can and has been interpreted in a variety of different ways. As Fine (1996) says, it is "no easier to characterize constructivism than it is to characterize realism or antirealism" (Fine 1996, p.233).

Despite Fine's qualms about the difficulty of characterising constructivism there is at least some agreement on the importance of recognising the difference between the implicit constructivist claims of the Strong Programme (Bloor 1976) and a stronger brand of constructivism endorsed by the likes of Latour and Woolgar (1979), Knorr-Cetina (1981), Pickering (1984) and Latour (1987). As Hacking says, in contrast to these explicitly constructivist studies, "constructionism does not seem to be so intimately involved in the Strong Programme as is commonly made out" (Hacking 1999, p.65). In what follows I want to back up this claim by showing how and why we must learn to properly distinguish between the macro-social explanations of the Strong Programme and the micro-social accounts of practice developed by the likes of Latour, Woolgar and Pickering. I will suggest that referring to both of these projects as "constructivist" only serves to confuse and blur the important differences between them.

## 5.1 The strong programme in the sociology of scientific knowledge

The decisive move in the development of modern sociology of science was the rejection of the Mertonian assumption that the job of the sociologist is to analyse the "cultural structure" rather than the substantive findings or content of science (Merton 1973, p.268). As long as this "arationality assumption" (Laudan 1977) remained it was possible for philosophers to distinguish their own work on the logical structure of science from the sociological investigation of scientific institutions and communities. This was enshrined in the distinction between the context of justification and the context of discovery, a distinction designed to protect the substantive content of science from sociological or psychological analysis. As is well known, Kuhn (1970) famously undermined this distinction by arguing that the transition from one paradigm to another is not simply a matter of applying shared rational criteria. For Kuhn, the fact that competing paradigms are 'incommensurable' means that "when paradigms enter, as they must, into a debate about paradigm choice their role is necessarily circular. Each group uses its own paradigm to argue in that paradigm's defense" (Kuhn 1970, p.74).

Kuhn's suggestion that there are no theory-neutral criteria available for assessing competing paradigms seemed to imply the need to invoke further factors in the explanation of how the transitions between paradigms occur. In particular, Kuhn's account suggests that, pace Mannheim and Laudan, *social* factors may after all be relevant to the content of science. This suggestion was first put into a programmatic form by Bloor (1976) who argued for a radical new approach to the sociology of science based on four tenets or maxims:

> 1. It would be causal, that is, concerned with the conditions which would bring about belief or states of knowledge. Naturally, there will be other

types of causes apart from social ones which will co-operate in bringing about belief.

2. It would be impartial with respect to truth and falsity, rationality or irrationality, success or failure. Both sides of these dichotomies will require explanation.

3. It would be symmetrical in its style of explanation. The same types of cause would explain, say, true and false beliefs.

4. It would be reflexive. In principle its patterns of explanation would have to be applicable to sociology itself

Bloor (1976, pp.4-5)

These four methodological principles are the basis of the 'strong programme' in the sociology of knowledge, an attempt to explain the structure *and* content of science from a sociological perspective.

Studies of science based on Bloor's four methodological principles appeared in the late 1970s and continued to emerge throughout the 1980s. A characteristic feature of these studies was the way in which they attempted to explain the resolution of scientific controversies in terms of the wider social interests of their participants. As Golinski (1998) argues, this concern with specifying the *macro-social* causes of scientific beliefs was one of the ways in which the strong programme went beyond Kuhn. However, as time went on there was a gradual move away from the attempt to explain scientific controversies in terms of *external* social factors. Thus, although Shapin and Schaffer (1985) attempted to show how the debate between Hobbes and Boyle over the status of the air pump was related to wider political and sociological concerns, other studies preferred to focus on more localised factors in the settlement of disputes and the production of knowledge (Pickering 1984; Collins 1985).

## 5.2 The social construction of facts: ethnographic studies of science

The shift away from the macro-social approach of the strong programme was partly a result of a gradual change in the locus of sociological investigation. As Fuller (1993) says:

> Bloor highlighted the extent to which scientific disputes are social struggles in symbolic disguise. However, once the sociologists started to conduct ethnographies of science as it was actually been done, they discovered a massive disparity between the words and deeds of scientists. The words as they appeared in journal

> articles, largely conformed to philosophical canons of rational methodology, but they also represented a highly idealized – if not downright misleading – picture of what took place in the lab. The lab work turned out to be quite chaotic and open-ended, even at the level of personal interests and group understanding of the ends of their research. (Fuller 1993, p.9)

The *locus classicus* of this ethnographical approach to the analysis of laboratory science is the seminal study of Latour and Woolgar (1979) in which the experimental and theoretical practice of scientists is examined *in situ* from the perspective a "naïve" sociological investigator. The purpose of this and other ethnographical studies (Knorr-Cetina 1981; Lynch 1984) is to show how knowledge is intimately connected to the site of its production.

Aside from differences in the way they carry out their investigations there is also an important difference between the strong programme and ethnographic studies in what they have to say about the status of science. In particular, proponents of the ethnographic approach generally argue for a far stronger constructivist thesis than the macro-social studies of the strong programme. As Kukla (2000) says:

> SSKists are often called constructivists on no more basis than that they regard scientific beliefs as having social causes. However, *some* SSKists – by no means all of them – have advanced a far more adventurous thesis. According to Latour, Woolgar, Knorr-Cetina, Collins and Pickering (*inter alia*), it's not only scientific *beliefs* that are socially constructed – it's scientific *facts*. (Kukla 2000, pp.8-9)

According to this stronger constructivist claim science is constructed in a much stronger sense than was previously thought. The strong programme suggests that the beliefs of scientists are 'constructed' from a matrix of social causes including personal interests and allegiance, but if the strong constructivist thesis is correct then reality itself is a product of the decisions and procedures scientists. Kukla (2000) has tried to capture the difference between these two claims by referring to the first as epistemic constructivism (because it concerns what we believe) and the latter as metaphysical constructivism (because it concerns what there is). Throughout this thesis I shall employ these terms.

### 6. Three varieties of naturalistic realism

We now have some idea of the kind of positions and arguments that confront the naturalistic realist. However, we have yet to give an account or an explanation of just what naturalistic realism is. It is time now to remedy that situation. At the risk of stating the obvious, naturalistic realism is the attempt to defend a realist account of science from a naturalistic perspective. From the previous chapter we know that there is an important difference between weak scientistic naturalism (the attempt to construct an account based on the methods of science) and strong scientistic naturalism (the attempt to construct an account based on the methods *and* the results of science). Given the discussion of that chapter one may be tempted to conclude that strong scientistic naturalists tend to be realists of one sort or another (e.g. Kitcher and Giere) whereas weak scientistic naturalists, i.e. Larry Laudan, are antirealists. However, this sort of conclusion is misleading in two important respects. Firstly, not all strong scientistic naturalists are realists because sociologists of science clearly rely on both the results and methods of science (i.e. sociology) in order to support antirealist conclusions (i.e. epistemic and metaphysical constructivism). Secondly, not all (naturalistic) realists are strong scientistic naturalists because, like Laudan, proponents of the explanationist defence invoke only the methods rather than the results of science in their defence of realism. Let me respond to these worries.

The omission of sociology of science can be defended on the grounds that the last chapter was primarily concerned with understanding naturalism as an attempt to carry out the traditional *normative* mission of philosophy of science. Once understood in this way it is not unreasonable to say that most strong scientistic naturalists, at least in the philosophy of science, are realists. However, in response to the second objection, we should grant that strong scientistic naturalism is not the only way to argue for realism from a naturalistic perspective. Thus, we must acknowledge that one can use the methods of science in order to argue for realism without necessarily having to appeal to the results of science. In order to capture this fact I will now outline three of the most prominent attempts to defend realism from a naturalistic perspective, two of which are based on using the methods *and* results of science (Kitcher 1993 and Giere 1988) and one that relies on methods only (Psillos 1999).

### 6.1 Giere's constructive realism

One of the chief aims of Giere's 'cognitive approach' to the philosophy of science (1985, 1988) is to show that naturalism is compatible with a modest version of scientific realism. *Pace* Cartwright (1983), Giere argues that the laws of physics "cannot tell lies about the world because they are not really statements about the world" (Giere 1988, p.90). Instead, he suggests that we think of laws and equations as defining a set of abstract theoretical models rather than representing real systems:

> We may even speak here of "truth." The interpreted equations are *true of* the corresponding model. But truth here has no *epistemological* significance. The equations truly describe the model because the model is defined as something that exactly satisfies the equations. (Giere 1988, p.79)

On this view, laws and equations can be said to be true but only in a trivial sense, it is their associated models that do the real representational work in science. However, this is not to say that models are *true of* real systems. As non-linguistic entities models are not even a candidate for truth or falsity. For Giere, it is a *theoretical hypothesis*, "a statement asserting some sort of relationship between a model and a designated real system" (Giere 1988, p.80), that is either true or false, not the model itself. These models are to be thought of as being *similar* to real systems rather than being true of them (or even 'isomorphic' to them as van Fraassen [1980] suggests). The role of theoretical hypotheses is simply to specify the respects and degrees of the similarity in question. Thus, for Giere, one can make a case for realism without presuming that all aspects of a model have counterparts in the world

Although the ultimate cause of Giere's commitment to realism may lie elsewhere it is clear that Laudan's pessimistic induction, van Fraassen's constructive empiricism, and social constructivism are major influences on the form that Giere's constructive realism takes. Laudan's influence, although rarely mentioned, is clear from the fact that Giere attempts to *test* his theory against the actual practice of science. In contrast, the influence of van Fraassen's arguments is much more clear. Indeed, aside from his disagreement with van Fraassen the realism-antirealism issue, Giere's account is very much like constructive empiricism in the way it attempts to understand the theory and practice of contemporary physics via the semantic approach to scientific theories.[20] Similarly, although he is "repelled by the idea that science is purely a social construct" (Giere 1988, p.xvi), Giere is prepared to accept that many aspects of science may indeed be 'socially constructed.' For example, Giere allows, "that the systems described by the various equations of motion are *socially* constructed entities. They have no reality beyond that given to them by the community of physicists" (Giere 1988,

[20] As Giere admits, his own position – 'constructive realism' – takes its name from the fact that it is intended to be a realist version of van Fraassen's constructive empiricism. See Giere 1988, p.93.

p.78). Further, Giere acknowledges that the "family resemblance" of models "is solely a matter to be decided by the judgments of the scientific community at the time. This is not to say that the collective judgments of scientists *determine* whether the resemblance is sufficient. This is one respect in which theories are not only constructed, but socially constructed as well" (Giere 1988, p.86).

Although he acknowledges their role in highlighting various important features of science, Giere ultimately rejects both van Fraassen's constructive empiricism and the social constructivist view of science. He does so for the familiar reason that they fail to provide an adequate explanation of science (Giere 1988, pp.128-131). In order to support this negative appraisal Giere introduces a study of nuclear physicists working in a national cyclotron facility that is explicitly based on the ethnographic approach of Latour and Woolgar (1979) and Knorr-Cetina (1981). Unsurprisingly, Giere concludes that constructive realism is, in this case at least, the best (and perhaps only) explanation of the behaviour of scientists:

> The only remotely plausible *scientific* account of what physicists are doing requires us, as students of the scientific enterprise, to invoke entities with roughly the properties physicists themselves ascribe to protons and neutrons. (Giere 1988, p.112)

So, it turns out that Giere's argument for scientific realism is just another version of the explanationist defence defended by Boyd (1983) and Putnam (1975). Having said this, there are of course important differences between Giere's argument and the standard formulation of the explanationist defence. Firstly, in Giere's argument the *explananda* concerns similarity rather than truth or reference and the *explanandum* is to account for how scientists talk and behave rather than the success of science.[21] Secondly, in contrast to original formulations of the explanationist defence, Giere is careful to restrict his conclusions to cases where he can actually provide concrete evidence for the superiority of a realist explanation (Giere 1988, p.112).

### 6.2 Kitcher's progressive realism

In *The Advancement of Science*, Kitcher (1993) presents us with one of the most sophisticated attempts to defend a naturalistic account of realism and progress in science. According to Kitcher, science consists of a multi-faceted set of practices the

---

[21] Giere does appeal to the success of science in his introduction when discusses the failures of social constructivism. See Giere 1988, p.4.

components of which improve as we progress through time. He suggests that by idealizing from the complexity of science we can identify the following components of the *individual* practices of scientists: scientific language, significant questions, accepted statements, explanatory patterns (or schemata), credibility judgments, paradigms of experimentation and observation, and exemplars of good and faulty scientific reasoning. The *consensus* practice of a particular scientific community consists of the components of individual practices that all members of the community share.

Progress in science results from improvement in consensus practices that in turn results from changes in individual practices caused by social interactions with colleagues and asocial interactions with nature. However, to say that the move from one consensus practice to another is progressive need not imply that we have improved every component of our practice. For Kitcher, there are several distinct ways in which a particular consensus practice can improve on its predecessor. The most important kinds of progress concern improvements made to three particular components of consensus practice: the language we use to describe phenomena, the explanatory patterns we use to explain phenomena, and the questions we ask.

Firstly, improvement in the language we use, or *conceptual progress*, "is made when we adjust the boundaries of our categories to conform to kinds and when we are able to provide more adequate specifications for our referents" (Kitcher 1993, pp.95-96). In Kitcher's view, we can only recognize the existence of conceptual progress by rejecting the presupposition that there should be a uniform mode of reference for all tokens of a single type. Instead, Kitcher argues that the reference potential of a scientific term may have a "heterogeneous reference potential" (ibid. p.101), by which he means that the mode of reference for a particular term token can be one of three types: the descriptive type, the baptismal type or the conformist type.

The second kind of progress that is fundamental to Kitcher's account concerns improvement in our explanatory patterns, or *explanatory progress*, which "consists in improving our view of the dependencies of phenomena" (Kitcher 1993, p.105). For Kitcher, science progresses not only by improving the reference potentials of key terms but also by being better able to fathom the "objective order of dependency in nature" (ibid. p.106). This order is instantiated in the explanatory schemata that scientists use to explain the way in which phenomena relate to each other. Changes in the way scientists view dependencies between phenomena can be seen in the way in which these explanatory schemata change through time. If these changes are made in such a way that they come closer to explaining actual dependencies in nature then we have made explanatory progress.

The third kind of progress that is fundamental to Kitcher's account concerns improvements made in the questions that are taken to be significant by a particular scientific community, which "arise against the background of the ordering of the phenomena captured in our explanatory schemata" (ibid. p.112). According to Kitcher, there are two main types of significant question that relate to explanatory schemata. *Application* questions "are generated from projects of finding particular instantiations of the available schemata" (ibid. p.112), and *presuppositional* questions "investigate the conditions that must obtain if the available schemata are to be instantiated" (ibid. p.113). *Erotetic* progress is made in science when we are able to pose application and presuppositional questions that are genuinely significant with respect to a certain set of explanatory schemata and that were not previously asked before.

Despite his insistence that it need not be interpreted as such (ibid. p.134), this account of progress is clearly intended as a defence of scientific realism. This is apparent in Kitcher's claim that the aim of science is to attain significant truth and is most obvious when he advertises the virtues of his account:

> It offers a broadly realist view of science: scientists find out things about a world that is independent of human cognition; they advance true statements, use concepts that conform to natural divisions, development schemata that capture objective dependencies. (Kitcher 1993, p.127)

Further evidence for Kitcher's commitment to realism can be found in the way he responds to some of the positions and arguments discussed in this chapter. For example, Kitcher shows how we can use his naturalized account to counter Laudan's pessimistic induction (ibid. pp.140-149), van Fraassen's constructive empiricism (ibid. pp.150-157) and the claim made by some sociologists of science that social interference rules out unbiased access to nature (ibid. pp.160-169). Like Giere, Kitcher's aim is to show that realism (in the form of his account of progress) provides a better explanation of science than these competing antirealist alternatives.

## 6.3 Psillos's (weak) naturalistic realism

Kitcher and Giere base their realist accounts of science on empirical information taken from cognitive psychology and evolutionary theory. In this sense their commitment to scientific realism is based on what we have been calling strong scientistic naturalism.

However, as I suggested above, not all naturalistic realists endorse such a strong commitment to the results of science. Thus, in his recent book, Psillos (1999) endorses a form of naturalistic realism that is based purely on an appeal to the methods of science:

> In its attempt to investigate the epistemic credentials of science, and in particular to answer the question of why scientific methodology is instrumentally reliable, a realist epistemology of science should employ no methods other than those used by scientists themselves (Psillos 1999, p.78)

Following Boyd (1983), Psillos argues that whether or not realism is the best explanation of science is an empirical matter to be decided by appealing to appropriate evidence.[22] Here Psillos suggests that when such evidence is considered it allows us to conclude that "there are no better explanations that the realist one" (Psillos 1999, p.97). However, unlike Kitcher and Giere, it is not Psillos's aim to show that a realist account of science must, *in addition*, be based on the results of a particular science. Like Boyd, Psillos has little or no interest in developing this stronger kind of naturalistic account.

Although he does not share their commitment to strong scientistic naturalism, Psillos's attempt to defend realism covers much of the same ground as the accounts of Kitcher and Giere. For example, he attempts to show how we can develop a realist account of reference and truth that will enable us to refute Laudan's pessimistic induction. Here Psillos shows that his commitment to naturalism is more than just a promissory note because he attempts to 'test' the empirical hypothesis of realism against detailed historical case studies (Psillos 1999, ch.6). Psillos also follows Kitcher and Giere in identifying van Fraassen's constructive empiricism as another major opponent of naturalistic realism. Here he argues that there are certain reasons for preferring a realist account of science to van Fraassen's empiricist alternative (ibid. ch.9). In fact, in terms of their identification of the major opponents of naturalistic realism, the only real difference between Psillos's account and the strong naturalistic accounts of Kitcher and Giere is that the former has very little to say about recent constructivist challenges to realism.

---

[22] As Rosenberg (1996) points outs out, Boyd (1983) was one of the realists to realise that there was no prospect of settling the issues that divide realists and antirealists on non-naturalistic grounds. See Rosenberg 1996, pp.5-6.

## 7. Conclusion

In this chapter we have seen that some philosophers of science attempt to use the 'naturalistic turn' in order to defend a realist account of science. In doing so they see their accounts as viable realist alternatives to three particular varieties of antirealism: Laudan's naturalistic antirealism, van Fraassen's constructive empiricism, and various 'constructivist' accounts provided by recent sociology of science. However, if naturalistic realism is to be preferred over these accounts then they need to show that the arguments advanced in favour of antirealism are not sound. In the next chapter I discuss the attempts by two naturalistic realists (Kitcher and Psillos) to show that Laudan's pessimistic induction is not sufficiently strong to support the conclusion that current theoretical terms (probably) do not refer.

# Chapter 3

# A Defence of Laudan's Pessimistic Induction: Part 1

## 1. Introduction

If the analysis of the previous chapter is correct then one of the most serious obstacles in the way of naturalistic realism is Laudan's pessimistic induction on the history of science. Indeed, it is no exaggeration to say that defeating the pessimistic induction has become almost a defining feature of modern-day scientific realism, naturalistic or otherwise. As Giere (1988) says:

> Many philosophers of science remain unconvinced by Laudan's arguments, but no realistic theory of science can be viable if it fails to account for the historical evidence Laudan presents. (Giere 1988, p.46)

Unfortunately, no matter how unconvinced realists may have been, attempts to refute the pessimistic induction have in one way or another failed to provide a convincing knockdown argument against it. In many ways this is a consequence of the way in which realists like Boyd (1983) shifted the realism-antirealism issue away from technical questions concerning reference and truth to questions concerning the historical necessity of a realist account of science. Laudan's claim that this account is refuted by the history of science has required the realist to engage in protracted disputes concerning the proper interpretation of examples. Because of this it has always been possible for one side or the other to question the methodological and historiographical assumptions behind their opponent's case studies. My aim in this chapter is to show that recent attempts to settle such issues in favour of realism by naturalists like Kitcher (1993, 2001) and Psillos (1999) are not successful.

In the following section I show that although realism has attracted many new opponents since its return in the 1970s this has not resulted in much change in the basic argument used to defend it. Most naturalistic realists in the philosophy of science endorse one or another form of the explanationist defence of scientific realism. In section 3, I discuss the two main ways in which the realist can reduce the number of counterexamples on Laudan's list: the 'wrong-sort-of-success' and 'maturity' strategies.

I argue that although these strategies can plausibly be used to dismiss some of Laudan's counterexamples, they cannot de used to dismiss them all. In section 4, I show that a purely causal theory of reference as advocated by Hardin and Rosenberg (1982) is not a good way of dealing with Laudan's more problematic cases and, in section 5, argue that a causal-descriptive approach like the one provided by Kitcher (1993) is likely to be more successful. However, in section 6, I discuss the work of Solomon (1995) who suggests that there is a crucial problem with Kitcher's attempt to analyse past theories in terms of working and presuppositional posits. In section 7, I follow a suggestion of Matheson (1996) and argue that even if the problems raised by Solomon (1995) could be resolved, Kitcher's response to Laudan generates a new version of the pessimistic induction. In section 8, I look at Psillos's recent attempt to respond to Laudan to see if it does any better. In section 9, I argue that although Psillos can solve some of the problems generated by Kitcher's account he nevertheless fails to conclusively refute the new version of the pessimistic induction. In section 10, I conclude that despite the efforts of Kitcher and Psillos, Laudan's pessimistic induction (albeit in an altered form) still constitutes a powerful argument against realist philosophies of science.

## 2. Scientific realism and the pessimistic induction

Laudan's (1981/1996) pessimistic induction is a response to the claim that "epistemological realism is an empirical hypothesis, grounded in, and to be authenticated by, its ability to explain the workings of science" (Laudan 1996a, p.107). At the vanguard of this move to defend realism on empirical grounds Laudan identified the work of Boyd (1973), Newton-Smith (1978), Putnam (1975), and Niiniluoto (1977). With the benefit of hindsight we can see this move toward naturalistic arguments for realism as being the result of two contributing factors. Firstly, there was the gradual development of the 'no miracles' argument for realism, the claim that it would be a miracle if mature scientific theories were successful but not true (or approximately true). As Psillos (1999, pp.72-77) has persuasively argued, this argument has its origins in Smart's 'no cosmic coincidences' argument (Smart 1963, p.39) and Maxwell's attempt to defend realism as the only reasonable explanation of science (Maxwell 1962, p.18). Secondly, there was the development of the causal theory of reference (Putnam 1975, Kripke 1980) according to which reference is fixed purely in terms of a baptismal event where a speaker enters into a certain causal relation with an object. The significance of this theory is that we can still be said to be referring to an object even though its description or properties may change. On this view, reference is a purely naturalistic, causal relation between speakers and the world.

If one combines the reasoning behind early formulations of the 'no miracles' argument and the causal theory of reference then one will arrive at what Psillos has called the 'Boyd-Putnam' argument for realism. This argument claims that the best scientific explanation of the empirical success of science is that the central theoretical terms of mature scientific theories genuinely refer, i.e. scientific theories are based on genuine causal relations between theories and the world. Of course, it is precisely this empirical claim that Laudan is out to rebut. The basis of his challenge is the identification of certain theories in the history of science that appear to undermine the realist's 'miracle' intuition that the empirical success of a scientific theory can only be explained by the fact that the central theoretical terms of the theory genuinely refer. Before continuing, let us remind ourselves of Laudan's reconstruction of the realist's argument:

(RS-1) The theories in the advanced or mature sciences are successful.

(RS-2) A theory whose central terms genuinely refer will be a successful theory.

(RS-3) If a theory is successful, we can reasonably infer that its central terms refer.

(RS-4) All the central terms in theories in the mature sciences do refer.

The sophistication of Laudan's attack lies in his attempt to undermine both the claim that success betokens reference (RS-3) and the claim that reference guarantees success (RS-2). In other words, *pace* Boyd and Putnam, Laudan thinks he has shown that genuine reference of central theoretical terms is neither necessary nor sufficient for the empirical success of science.

In responding to Laudan's attack on convergent realism, most realists have been concerned with the argument against the inference from success to reference (RS-3) rather than that against the inference from reference to success (RS-2). This is arguably a result of two factors. Firstly, it is crucial for the realist's purposes that we are able to infer from the success of a theory that its central theoretical terms genuinely refer and as a consequence that it is approximately true. Although the reverse inference is just as important for securing the connection between success and reference, it is only the inference from success to truth (via reference) that can get the realist's project off the ground. Secondly, it is generally agreed that Laudan's many examples of successful-yet-false theories present a much stronger case against RS-3 than the relatively few examples of true-but-unsuccessful theories presented against RS-2. As a consequence, it is not uncommon for responses to Laudan to focus purely on the case he makes against RS-3. Indeed, in most cases, it is the argument against RS-3 that people have in mind when they talk about the pessimistic induction.

In order to appreciate the various realist responses to the pessimistic induction it is instructive to see how they attempt to weaken one or more of the premises in the following reconstruction of Laudan's argument:

**The pessimistic induction – version 1**

1. If a theory is successful, we can reasonably infer that its central terms refer.

2. Most current scientific theories are successful.

3. We can reasonably infer that the central terms of most current scientific theories refer.

4. If the central terms of most current scientific theories refer, then the central terms of most past scientific theories did not refer since the reference of these terms was fixed by faulty theoretical descriptions.

5. Many of these past scientific theories were successful.

6. If a theory is successful, we cannot reasonably infer that its central terms refer.

This is the pessimistic induction in its strongest and most persuasive form whereby a *reductio ad absurdum* is carried out on the realist's inference from success to reference.[1] As I will now show, the realist responds to this argument by deploying a number of strategies that attempt to alter one or more of its premises in the hope that (eventually) the conclusion (statement 6) will no longer inductively follow from the first five premises.

[1] My reconstruction largely follows that of Peter Lewis (2001) whose argument against the pessimistic induction will be considered in the following chapter.

### 3. Reducing the realist's workload: redefining success and maturity

In responding to Laudan (1996a), realists have typically chosen to take up Laudan's historical challenge by providing a realistic interpretation of the examples used to undermine the alleged connection between reference and success.[2] However, this strategy is not strictly necessary for all of the examples on Laudan's list. As Matheson (1996) points out, it is perfectly possible for realists to object that many of the examples of successful-yet-false theories on Laudan's list exhibit the 'wrong sort' of success:

> In many cases, realists can claim that the putatively successful theories listed by pessimists like Laudan were not successful in the right way, because they failed to succeed to a certain degree, because their success was not prolonged, or because they never reached an appropriate rate of success. (Matheson 1996, p.469)

Of course, if this strategy is to be anything other than an ad hoc attempt to rule out problematic examples then the realist had better tell us what it is that makes a theory *genuinely* successful. Here realists have been almost unanimous in suggesting that a theory is genuinely successful only when it generates novel predictions (McMullin 1987; Worrall 1982, 1994; Psillos 1999, pp.105-107). As Psillos (1999) says, "the notion of empirical success should be *more* rigorous than simply getting the facts right, or telling a story that fits the facts. For any theory (and for that matter, any wild speculation) can be made to fit the facts – and hence to be successful – by simply 'writing' the right kind of empirical consequences into it. The notion of empirical success that realists are happy with is such that it includes the generation of novel predictions which are in principle testable" (Psillos 1999, p.105).

By employing the 'wrong-sort-of-success' strategy realists have persuasively argued that at least some of the examples on Laudan's list were not genuinely successful because they did not generate novel predictions of phenomena, e.g. the crystalline spheres theory, the theory of circular inertia, contact-action gravitational theories, etc. In terms of the argument outlined in the previous section this sort of response results in the following modified formulation of Laudan's argument, where changes to that argument are indicated in italics:

**The pessimistic induction – version 2**

1. If a theory is *genuinely* successful, we can reasonably infer that its central terms refer.

2. Most current scientific theories are *genuinely* successful.

3. We can reasonably infer that the central terms of most current scientific theories refer.

4. If the central terms of current scientific theories refer, then the central terms of most past scientific theories did not refer since the reference of these terms was fixed by faulty theoretical descriptions.

5. *Some* of these past scientific theories were *genuinely* successful.

6. If a theory is *genuinely* successful, we cannot reasonably infer that its central terms refer.

Here we can see that the 'wrong-sort-of-success' strategy results in two important changes to the original argument. Firstly, in premises 1, 2, 5 and 6 we have the qualification that Laudan's argument must be restricted to *genuinely* successful scientific theories. Secondly, and as a corollary, premise 5 has changed from the claim that most past scientific theories were successful to the weaker claim that only *some* of these theories were genuinely successful.

So, by deploying the 'wrong-sort-success' strategy the realist can claim to have at least some grounds for questioning the inductive evidence from which Laudan generates his pessimistic conclusion. However, not all of the theories on Laudan's list can be dealt with in this fashion. Fortunately, there is another way in which the realist can tackle at least some of the remaining historical counterexamples. This is to say that although certain theories may have been both false and genuinely successful they may not be *mature* and are therefore not part of the realist's explanandum:

> Another way to reduce the size of Laudan's list is to suggest that *not all* past theoretical conceptualisations of domains of inquiry should be taken seriously.

[2] Although there is generally some attempt made to respond to the examples used to block the 'reference to success' move (RS-3), the majority of philosophical ink has been spilled on realist interpretations of the examples that seem to threaten the 'success to reference' (RS-2) move.

> Realists require that Laudan's list should only *mature* theories; that is, theories which have passed the 'take-off point' (Boyd) of a specific discipline. (Psillos 1999, p.107)

If this strategy is to avoid the charge that it is simply an ad hoc attempt to rule out problematic examples then the realist needs to give us an account of what 'maturity' consists in when applied to theories. Here is what Psillos (1999) has to say on this matter:

> This 'take-off point' can be characterized by the presence of a body of well-entrenched background beliefs about the domain of inquiry which, in effect, delineate the boundaries of that domain, inform theoretical research and constrain the proposal of theories and hypotheses. This corpus of beliefs gives the broad identity to the discipline by being, normally, the common ground that rival theories of the phenomena under investigation share. It is an empirical matter to find out when a discipline reaches the 'take-off point', but for most disciplines there is such a point (or, rather a period). (Psillos 1999, pp.107-108)

For Psillos, this empirical requirement for the maturity of theories allows us to reject yet more of the examples on Laudan's list, e.g. the humoral theory of medicine and the effluvial theory of static electricity.

Again, the 'maturity' strategy is best seen in terms of how it seeks to weaken the original formulation of Laudan's argument:

**The pessimistic induction – version 3**

1. If a *mature* scientific theory is genuinely successful, we can reasonably infer that its central terms refer.

2. Most current *mature* scientific theories are genuinely successful.

3. We can reasonably infer that the central terms of current *mature* scientific theories refer.

4. If the central terms of most current *mature* scientific theories refer, then the central terms of most past scientific theories did not refer since the reference of these terms was fixed by faulty theoretical descriptions.

5. A *few* of these past *mature* scientific theories were genuinely successful.

6. If a *mature* scientific theory is genuinely successful, we cannot reasonably infer that its central terms refer.

So, when combined with the modifications from the 'wrong-sort-of- success' strategy, the claim that we must restrict Laudan's argument to mature scientific theories results in two important modifications. Firstly, in all five premises and also the conclusion we have the qualification that only *mature* scientific theories are under consideration. Secondly, in premise 5 we have moved from the claim that some past scientific theories were genuinely successful to the weaker claim that only a *few* past *mature* scientific theories were genuinely successful.

The realist would like to be able to say is that this double attack on premise 5 weakens it to the point where it can no longer function as the antecedent in the conclusion. If we define success and maturity in such a way that *none* of Laudan's counterexamples are legitimate then this is precisely what we will be able to conclude. However, it is hard to see how such definitional legerdemain could avoid the charge of being a purely ad hoc attempt to rule out problematic counterexamples. Of course, we must allow the realist at least some leeway in responding to the pessimistic induction by granting that it is not entirely up to Laudan to define the terms of the debate. In other words, even the most committed antirealist should accept the realist claim that "Laudan has overstated his case against antirealism" (Psillos 1999, p.104). However, from an antirealist perspective, it is also arguable that we must defend the pessimistic induction against the kind of realist trickery that would see *all* of his counterexamples crumble in the face of the 'wrong-sort-of-success' and 'maturity' strategies. How then can we decide which of Laudan's counterexamples are legitimate and which are not?

Fortunately, there is no need to go into any great detail in answering this question because there is a certain amount of consensus on which of Laudan's examples pose *genuine* difficulties for realism (i.e. difficulties that cannot be answered by either the 'wrong-sort-of-success' or 'not-yet-mature' strategies). Thus most commentators on the pessimistic induction now suggest that Laudan's case primarily rests on two examples of successful past theories whose central terms failed to refer: the caloric theory of heat and nineteenth-century optical theories based on the luminiferous ether. Although it is by no means certain that these are the *only* counterexamples that survive the two strategies discussed in this section, for the moment let us assume that they are. How serious a threat do these theories pose for realism?

## 4. Hardin and Rosenberg

If we accept that the 'wrong-sort-of-success' and 'maturity' strategies allow the realist to drastically reduce the number of Laudan's counterexamples then it seems we have a rather easy way of dismissing his pessimistic conclusion. This conclusion relies on having a sufficiently large number of examples where a scientific theory is successful yet its central theoretical terms do not refer. However, if the discussion of the previous section is correct then Laudan only has two *genuine* examples. So, it seems that Laudan does not have anywhere near enough evidence to support his pessimistic conclusion. Does this show that the pessimistic induction fails to work as an argument against the realist's inference from the success of a theory to the claim that its central theoretical terms genuinely refer?

Although there are an increasing number of realists who now endorse something like this response to Laudan, initial responses to the pessimistic induction tacitly assumed that realism could only be defended by showing that *all* of Laudan's counterexamples were somehow compatible with the inference from success to reference. For example, one of the earliest attempts to respond to the pessimistic induction by Hardin and Rosenberg (1982) claims that even after we have deployed the 'maturity' strategy:

> We are, nonetheless, left with several theories. Perhaps the most powerful and fertile of these were the ones which postulated a medium for the transmission of light and, later, electromagnetism. By any reasonable standard, such ether theories were post take-off. Furthermore, the non-existence of the ether is generally held to have been established by the negative result of the Michelson-Morley experiment. Here, if anywhere, Laudan should be able to make his case: if these successful theories failed to secure reference, then securing reference would seem to be irrelevant to success. (Hardin and Rosenberg 1982, p.611)

So, Hardin and Rosenberg accept that if genuinely successful and mature scientific theories based on the demonstrably non-existent ether fail to secure reference then Laudan has won.

However, it is at this point that Hardin and Rosenberg introduce their preferred strategy for dealing with such cases. This is to say that with the benefit of hindsight we can see that when past scientists used the theoretical term 'ether' they were actually referring to the electromagnetic field. On this *causal* view of reference, the fact that a theoretical term is based on a mistaken set of theoretical descriptions does not prevent that term from genuinely referring. We can sum this approach by saying that it results in the following revised form of Laudan's argument:

| **The pessimistic induction – version 4** |
|---|

1. If a mature scientific theory is genuinely successful, we can reasonably infer that its central terms refer.

2. Most current mature scientific theories are genuinely successful.

3. We can reasonably infer that the central terms of current mature scientific theories refer.

4. If the central terms of most current mature scientific theories refer, then *the central terms of most past scientific theories did refer since the reference of these terms was fixed causally and not by faulty theoretical descriptions.*

5. A few of these past mature scientific theories were genuinely successful.

6. If a mature scientific theory is genuinely successful, we *can* reasonably infer that its central terms refer.

So, the adoption of a purely causal theory of reference has two quite drastic consequences for the pessimistic induction. Firstly, it enables us to substantially change the central premise of Laudan's argument from the claim that the central terms of past scientific theories did not refer to the claim that they did refer. Secondly, and most importantly, it allows the realist to show that once the switch is made from a descriptive to a causal theory of reference it is no longer possible to generate the pessimistic conclusion that Laudan requires for his *reductio* of premise 1.

The move to a purely causal view of reference seems to be just what the realist needs in order to block the pessimistic induction. When combined with the 'wrong-sort-of-success' and 'maturity' strategies, this move promises to show that the central theoretical terms of scientific theories in Laudan's remaining counterexamples actually referred despite the fact that they were attached to largely false theoretical descriptions. Unfortunately, the problem with this sort of reply to Laudan's more problematic counterexamples is that, yet again, it seems to be a purely ad hoc attempt to preserve the connection between reference and success. If the realist is allowed to "adopt that account of reference which severs it from the detailed beliefs of the theorist" (Hardin and Rosenberg 1982, p.611), what is to stop us from seeing almost *any* theory one could care to name as essentially referring to entities whose existence we now countenance? The problem is that when pushed to its limit a purely causal theory of reference seems to make reference a trivial feature of scientific practice. No matter how wrong our theories are it will almost always be the case that we are somehow managing to refer. Perhaps some realists could accept this quasi-mystical view of our ability to refer but it is hard to see how it can do justice to the history of science.

## 5. Kitcher's naturalistic realism

Kitcher (1993) uses the account of progress discussed in chapter 2 to present a highly sophisticated response to Laudan's pessimistic induction. Kitcher begins by objecting to Laudan's claim that the realist must be a holist about theory confirmation. According to this holistic view, if a theory is empirically successful then this must be taken as confirmation of the *whole* theory, not as confirming some theoretical components rather than others:

> Part of what separates the realist from the positivist is the former's belief that the evidence for a theory is evidence for *everything* the theory asserts. (Laudan 1996a, p.116)

For Laudan, any attempt to slough off the offending non-referring theoretical parts of the theory will undermine the holistic conception of theory confirmation the realist relies on elsewhere. In particular, Laudan believes that the realist needs the holistic conception when confronted with empiricist attempts to restrict confirmation to the observable consequences of a theory only. Without the holistic conception the realist would seem to have no way of justifying his belief that success betokens the truth of unobservable aspects of the theory.

Putting aside for one moment the issue of whether or not realism is impossible without a holistic account of theory confirmation there still remains the important issue of whether or not it is possible to identify the constituents of a theory that are responsible for its success. Kitcher (1993) suggests that it is Laudan's obsession with holism that prevents him from seeing that different parts of a theory contribute differentially to its success. He introduces this argument by way of an analogy:

> Laudan's broad-brush way of discussing whole theories is reminiscent of the following "confutation" of those who think that successful basketball teams are typically those with tall players. We trot out examples of successful teams on which there is one diminutive person. It is, of course, important not to disclose the fact that this person has little or nothing to do with the team's success. Similarly, it is not enough to conceive a theory as a set of statements and distribute the success of the whole uniformly over the parts. One has to see how the statements are used. (Kitcher 1993, p.143)

Kitcher goes on to show how we might usefully apply this idea to one on the most important examples on Laudan's list – ether theories in nineteenth-century physics. He focuses on Fresnel's prediction of the Poisson bright spot, a successful example of scientific theorizing based on background theories whose central terms referred to a non-existent propagating medium. Without denying the obvious role that the ether played in nineteenth century physical theories Kitcher argues that, *pace* Laudan, Fresnel did manage to produce some important new truths about wave propagation and the behaviour of light. In order to do this he deploys the two most potent resources in his philosophical arsenal, the accounts of explanatory and conceptual progress in science.

Kitcher's first aim in responding to Laudan is to establish the precise role of the ether in Fresnel's explanatory schemata. Here Kitcher claims that, although the ether was *presupposed* by Fresnel's attempts to explain phenomena like interference and abstraction, it was never used in the explanatory schemata themselves. The importance of this point is that it allows Kitcher to drive a wedge between the subsequent rejection of the ether and the apparent success of the theories that presupposed it:

> All the successes of the schema can be preserved, even if the belief and the presupposition that it brings in its train are abandoned. That, to a first approximation, is what occurred in the subsequent history of wave optics. (Kitcher 1993, p.145)

Here in a nutshell is Kitcher's general strategy for responding to the examples on Laudan's list. The non-referring terms featuring in these examples are *presuppositional* rather than *working* posits and as such do not contribute to the success of the theory in question. In other words, only the working posits of a theory contribute to its empirical success.

The suggestion that the ether was a dispensable presupposition of Fresnel's theoretical practice is a good first step in deflecting Laudan's claim that apparently successful theories can contain non-referring central theoretical terms. However, as Kitcher recognizes, the realist also needs to respond to the claim that the presuppositional role of the ether was so strong that it permeated Fresnel's entire vocabulary, thereby making *all* of the terms in his theories non-referential. Kitcher's way of dealing with this possibility is to distinguish between two different ways in which we can fix the reference of a term-type like 'light wave':

1. *The Baptismal (or causal) Mode of Reference* - the dominant intention of the speaker is to talk about light, and the wavelike propagation of light.

2. *The Descriptive Mode of Reference* - the dominant intention of the speaker is to explicitly fix the reference of 'light wave' through a definite description based on background theories.

Applying these two modes of reference to the Fresnel case we can immediately see the value to Kitcher of adopting a causal-descriptive theory of reference. The claim that 'the ether' permeates all theoretical discourse (thereby making it non-referential) implicitly assumes that all term-tokens have their reference fixed *descriptively*, i.e. through definite descriptions based on a background theory that specifies the ether as a propagating medium. However, if there is also the possibility of fixing reference *causally*, i.e. talking about light without assuming anything about the propagating medium, then it seems perfectly possible that at least some of what Fresnel said about light (in particular the bits we would now accept) did genuinely refer. For Kitcher, it is only by acknowledging the existence of *both* of these modes of reference in Fresnel's work that we can reach an adequate understanding of his ability to refer to and explain the properties of light despite, not because of, his belief in a non-existent ether.

If we combine Kitcher's presuppositional vs. working posit distinction with his causal-descriptive theory of reference we can make the following alterations to the pessimistic induction:

**The pessimistic induction – version 5**

1. If a mature scientific theory is genuinely successful, *we can reasonably infer that its working posits refer and its presuppositional posits do not.*

2. Most current mature scientific theories are genuinely successful.

3. *We can reasonably infer that the working posits of current mature scientific theories refer and its presuppositional posits do not.*

4. If the working posits of most current mature scientific theories refer and its presuppositional posits do not, then *the working posits of most past scientific theories (sometimes) refer (i.e. when the reference of these terms is fixed causally rather than descriptively) and its presuppositional posits do not.*

5. A *few* of these past mature scientific theories were genuinely successful.

6. If a mature scientific theory is genuinely successful, *we can reasonably infer that its working posits (sometimes) refer and its presuppositional posits do not.*

Here we can see just how far Kitcher takes us away from Laudan's original argument. Firstly, in premises 1, 3, 4 and 5, we have the claim that the realist is only committed to a connection between success and the working posits of a theory. Secondly, in premise 4, we have the claim that such working posits refer only when reference is fixed causally as opposed to descriptively. Again, the significance of these adjustments is that they allow to Kitcher to block Laudan's pessimistic conclusion.

According to Kitcher, the value of his account of progress is that it allows us to avoid two equally implausible views of the history of science:

> I am not engaged in the whig enterprise of forgetting or even downplaying Fresnel's beliefs about the ether. Those should be given their due but should not be seen as some sort of creeping rot that invades everywhere. Laudan's penchant for tarring past theories as a whole with local misconceptions seems to me as unbalanced as neglecting every aspect of the past that does not fit within contemporary views of nature. (Kitcher 1993, p.148)

So, Kitcher's main argument for his realist account of progress is essentially another version of the explanationist defence, the claim that realism provides the best explanation of the success of science. However, unlike previous formulations of this argument, Kitcher claims that realist accounts that reject "every aspect of the past that does not fit within contemporary views of nature" (ibid. p.148) are just as bad as Laudan's holistic habit of saying that theories failed to refer simply because one of its terms is now though to be suspect. In other words, Kitcher thinks he is able to show why the purely causal view of reference adopted by Hardin and Rosenberg *and* the purely descriptive view of reference adopted by Laudan are *both* equally unacceptable attempts to explain relevant features of the history of science. We are invited to conclude that only Kitcher's account in terms of heterogeneous reference potentials and explanatory schemata can provide such an explanation.

## 6. Solomon's historical critique

Despite his claim to have provided an adequate philosophical explanation for the "mixed record of success and failure" (ibid. p.102) in the history of science, there are some who insist that Kitcher's account of the ether as a presuppositional posit rides roughshod over the historical details of nineteenth-century physics:

> For Maxwell, and other physicists at the time, the ether was a posit that was both *historically and psychologically* indispensable. His theory can be described so that the ether is *logically* dispensable – an idle additional presupposition – but that would be to rewrite history into a Legend that did not happen and probably could not have happened. (Solomon 1995, p.210)

Solomon's point here is that, despite his claims to the contrary, Kitcher is engaged in the whiggish attempt to project his own understanding of what was important about nineteenth century ether theories (i.e. which terms were working posits and which were presuppositional posits) back onto the history in order to reconstruct them to suit his own realist agenda.

Can Kitcher avoid the charge that the presuppositional-working posit distinction is a whiggish reconstruction of the role of the ether in nineteenth-century physics? Solomon (1995) thinks not. Although she accepts that it is perfectly possible to rewrite history in a way that makes it amenable to Kitcher's realist analysis she claims that establishing the *logical dispensability* of non-referring theoretical terms like the ether misses the point of Laudan's critique:

> Laudan's observation – that central claims in a successful theory can be abandoned in successor work – still stands, once the central claims are identified on the basis of their historical and psychological indispensability, rather than reconstructed logical indispensability. (Solomon 1995, p.210)

So, according to Solomon, Kitcher's strategy of establishing the logical dispensability of non-referring central theoretical terms does not provide a sufficiently robust interpretation of the examples on Laudan's list. As a consequence, Solomon demands that Kitcher show the historical and/or psychological dispensability (or idleness) of non-referring terms in the generation of successful theories and practices. In the case of ether theories Solomon takes this task to be all but impossible so, at least when interpreted in this way, Laudan's counterexample is immune to Kitcher's response.

Although Solomon is surely right to point out the whiggism inherent in Kitcher's response to Laudan, Kitcher might respond by making two points. Firstly, and most obviously, Kitcher's major historical example concerns the role of the ether in Fresnel's prediction of the Poisson bright spot, *not* the role of the ether in the development of Maxwell's theory of electrodynamics. Indeed, in response to Laudan's claim that for Maxwell "the aether was better confirmed than any other theoretical entity in natural philosophy" (Laudan 1984, p.14), Kitcher is quite prepared to bite the bullet and say that Maxwell was just plain wrong about this particular feature of his theoretical practice.

Secondly, it could be argued that Solomon's argument rests on a misconception of what Kitcher is trying to do. Kitcher is primarily concerned with defending the intuitive connection that is taken by realists to exist between success and reference. Laudan's examples seem to show that reference is neither necessary nor sufficient for the success of scientific theories. Thus, in order to restore the connection between reference and success, Kitcher needs to show how the non-referring central terms featuring in Laudan's examples do not contribute to the success of the theories in which they occur. As we have already seen, Kitcher does this by establishing the logical dispensability of central non-referring terms, e.g. the ether. In response, Solomon claims that in addition to showing the logical dispensability of a particular non-referring term Kitcher must also show its historical and psychological dispensability. But why must

Kitcher do this? Isn't it enough to show that non-referring terms like the ether can be removed from the historical picture without altering the predictive power and success of the theories that presupposed it?

In order to answer these questions we must first distinguish between two senses in which a term might be historically and/or psychologically indispensable. The first, trivial sense in which a term might be historically and/or psychologically indispensable is that, as a matter of historical and/or psychological fact, the scientists in the episodes in question used it. Clearly, when Solomon demands that Kitcher demonstrate the historical and/or psychological dispensability of terms like the ether it cannot be purely this sense of indispensability that she has in mind. If it were then it wouldn't even be worth Kitcher attempting to analyse the examples on Laudan's list. So it must be another sense of indispensability that is driving Solomon's argument and indeed, on closer inspection, this is precisely what we find:

> Unless he had posited an electromagnetic ether by analogy with continuum fluid dynamics, Maxwell would not have developed his theory at all. (Solomon 1995, p.210)

This second, more significant sense of indispensability involves the counterfactual claim that the ether played an indispensable historical and psychological role in the generation of Maxwell's theories, without the former we could not have had the latter. Note that this is not the trivial claim that we cannot give a full historical account of Maxwell's beliefs and theoretical practice if we do not include reference to the ether but rather that it played an essential role in his ability to produce empirically successful physical theories. It is this second counterfactual sense of indispensability that threatens Kitcher's analysis because, if true, it threatens to make the ether *indirectly* responsible for the success of theories it helps to generate.

Is it possible for Kitcher to demonstrate that the ether played no essential role in the generation of successful empirical theories? At first glance it would appear not; for looking at a case like Maxwell's it looks, as Solomon suggests, like the ether did indeed have some kind of essential historical *and* psychological role to play in the generation of his physical theories. However, it might be possible for Kitcher to claim that, although the ether is indispensable in any explanation of the *generation* of physical theories, it is dispensable when it comes to explaining the *success* of such theories. That

Kitcher might endorse something like this response to Solomon's critique is suggested in his most recent thoughts on this problem:

> Hindsight is well known to be wonderful in clarifying vision. Contemporary thinkers find it much easier to pick out the working posits and the idle wheels of the wave theory of light. One might even wonder if it's possible for practitioners to draw this distinction, so that, when we turn to our own theories, we'll find ourselves unable to identify those parts for which we can justifiably claim truth on the basis of success. Yet, even though Fresnel *didn't* make the distinction between those parts of his theory that were implicated in his successes and those that were not, it's far from obvious that he *couldn't* have done so. (Kitcher 2001, pp.18-19)

It is hard to know what we should make of this claim. Solomon argues that, as a matter of fact, the ether was historically and psychologically indispensable in the success of nineteenth century physical theories. In response, Kitcher tells us that although theorists like Fresnel failed to see that the ether did not contribute to the success of his theories, there is a possible world in which he might have done so. How is this counterfactual possibility supposed to help Kitcher's argument for the dispensability of the ether? How can the fact that I might have thought differently be used to justify the claim that I should have thought differently?

The problem with Kitcher's response is that a counterfactual claim about the dispensability of the ether in an alternative *possible* world is simply another way of saying that the ether was logically but not historically or psychologically dispensable in this *actual* world. Of course, Kitcher might say that the indispensability of the ether in this sense simply amounts to the trivial claim that any historical account of nineteenth century physical theories must include some reference to the ether. According to this suggestion, what is really at stake is the claim that the ether was historically and psychologically indispensable when explaining the success of these theories. Indeed, this must be what Kitcher means when he claims that it is far from obvious that Fresnel could have drawn a distinction between the ether as a presuppositional posit and the genuinely referring working posits of his theories. But what are the grounds for this claim other than the question begging idea that the ether was not responsible for the success of Fresnel's theories? The problem with Kitcher's possible worlds approach to the question of the dispensability of the ether is that it takes us round in a rather small and arguably vicious circle.

Perhaps there is a way in which Kitcher can show that the ether was not responsible for the successes of nineteenth century physical theories although it seems clear that counterfactual claims are not a good way of achieving this aim. Even if this were possible it might lead us to wonder just how much historical reconstruction is necessary in order to resurrect the connection between reference and success. However, even if it can avoid the charge of being ad hoc, there are still further difficulties facing Kitcher's account. Solomon's critique relies on the counterfactual claim that the ether was a *necessary presupposition* of Maxwell's theoretical practice. In this sense, she disagrees with Kitcher over the historical and/or psychological dispensability of the ether rather than over its status as a presupposition. As we will now see, there are some who even question Kitcher's claim that the ether functioned in a purely presuppositional role.

### 7. Matheson and the new pessimistic induction

Drawing on the historical work of Buchwald (1985, 1989), Matheson (1996) argues that according to Kitcher's own definitions the ether functioned as *both* a presuppositional posit and a working posit in nineteenth-century physical theories:

> Granted, some of these uses of the ether were failures but some (e.g. Fresnel on dispersion) were qualitatively successful. Therefore, Kitcher is incorrect when he says that the ether was explanatorily idle in every successful nineteenth century treatment of optical and electromagnetic phenomena. (Matheson 1996, p.472)

If Matheson is correct, the problem for Kitcher is that he will no longer be able to use the presuppositional-working posit distinction to show that the ether does not function in the explanatory schemata of nineteenth century physical theories. In other words, not only will he not be able to show that the ether is historically and/or psychologically dispensable, he won't even be able to show that it is *logically dispensable* either. The danger of this latter possibility being that Kitcher will again lose the desired connection between genuine reference and success.

So it seems that Kitcher's analysis of the ether as a presuppositional posit does not do justice to the historical details. How should we interpret this result? One possibility would be to respond to Matheson by pursuing the same kind of strategy that I suggested Kitcher might use to respond to Solomon, i.e. the use of the discovery-justification distinction to show that the success of a theory is attributable purely to its working posits. A second possibility would be to respond to Matheson either by denying that the ether functions as a working posit in the cases he alludes to, or alternatively claim that in any case in which it does function as a working posit the theory in question turned out to be empirically unsuccessful. Now it should be immediately apparent that the first sort of response is not available because Matheson's point is precisely that, on Kitcher's own terms, the ether was sometimes used as a working posit. This seems to leave Kitcher with the second response, the denial of Matheson's claim that the ether functioned as a working posit in successful theories. Is this a good way for Kitcher to respond to Matheson?

The first thing to say here is that only a detailed investigation of the examples alluded to by Matheson could settle the issue of whether or not the ether functions as a working posit in such cases. However, prior to such an investigation, there are two worries that might make us question the value of carrying it out, at least with respect to settling the issue in favour of Kitcher's account. Firstly, if the only way of preserving the presuppositional-working distinction is by modifying it to fit the cases alluded to by Matheson, how can we avoid the charge that the subsequent account will be purely ad hoc with respect to its explanandum? How do we respond to the suspicion that we are generating philosophical epicycles rather than genuine explanations? Secondly, even if we could reinterpret Matheson's examples in accordance with Kitcher's original formulation of the presuppositional-working distinction, what use would it be to us now?

Matheson (1996) argues that the problem with Kitcher's reconstructive tendencies is that they prevent us from distinguishing between presuppositional and working posits in contemporary scientific theories:

> Where we believe ourselves to have explained many things about the world via the properties of quarks, future Kitcherians might center these various explanations on entities which differ both from our current conceptions of quarks and from each other. Such a reconstruction may reassure future Kitcherians about the status of their science, but it should hardly make us feel good about ours. Although we can say of ourselves that we are talking about something, we can't say that we know what we are talking about. (Matheson 1996, p.473)

So, according to Matheson, even if it is post hoc as opposed to ad hoc, the problem with Kitcher's analysis is that, like Maxwell, we will never be aware of which constituents of our mature scientific theories genuinely refer (working posits) and which do not (presuppositional posits).[3] In fact, if we follow this chain of reasoning to its logical conclusion then things get a whole lot worse for Kitcher. Let me explain why.

Matheson claims that a reconstruction of past theories in terms of genuinely referring working posits and non-referring presuppositional posits is only available to future Kitcherians. For the sake of argument let us grant this point and say that the future Kitcherian's reconstruction at time $t + 1$ is of no use in helping us to decide which of our theoretical terms genuinely refer and which do not at the present time $t$. However, given this argument it is unclear why "such a reconstruction may reassure future Kitcherians about the status of their science" at time $t + 1$. After all, if our future Kitcherian is attempting to use this reconstruction to rebut a future Laudanian charge that *they* can have no confidence in the truth of *their* current scientific theories then Matheson's argument will be just as applicable at time $t + 1$ as it is to time $t$. So, pace Matheson, we will have to conclude that future Kitcherians are in the same epistemological boat as present-day Kitcherians for they are just as reliant on the judgments of their philosophical descendents as we are. Just as we can't say that we know what we are talking about at time $t$, neither will future Kitcherians know what they are talking about at time $t + 1$. It seems then that only future Kitcherians at time $t + 2$ will be able to tell which terms refer and which do not at time $t + 1$, but of course their judgments are equally prone to revision by future Kitcherians at time $t + 3$, and so on... *ad infinitum.*

So, things are a lot worse for Kitcher than we might have expected. The attempt to impose the working-presuppositional distinction on successful theories in the history of science in order to preserve the connection between reference and success has the rather curious result that although we can say we are talking about something we do not know what it is that we are talking about. At this point, Kitcher would surely point out that Matheson's argument (and my extension of it) relies on a hidden premise, namely that there is no important difference between present and past theories in terms of ability to accurately refer. Of course, the realist would like to be able to say that through the course of time our theories improve in terms of picking out natural kinds and that we also improve in our ability to know which of our theoretical terms genuinely refer and which do not. Thus, in response Matheson's argument, the realist can claim that

[3] A very similar claim is made by Solomon (2001): "according to Kitcher's views about reference and presuppositional posits, it cannot be concluded that our currently successful theories are substantially true or referential *in anything like the ways that we state and understand them*" (Solomon 2001, p.38)

although future Kitcherians will be better placed to see which of our theoretical terms refer and which do not, this is simply a reflection of the fact that our ability to pick out natural kinds improves through time. Similarly, in response to my extension of Matheson's argument, the realist can say that although every future generation is prone to the reconstructive judgments of its successors this is simply a result of two salient features of science: the acknowledgement that our theories are only approximately true and a belief in the progress of science.

Perhaps then Kitcher can avoid the charge that his account of working and presuppositional posits results in the bizarre conclusion that scientists past, present and future do not know what they are talking about. If theories improve through time then something like Matheson's futuristic scenario is precisely what we should expect. But, of course, the problem with this sort of reply to Matheson is that the realist is not entitled to assume that theories improve through time given that this is precisely what they are trying to show. In other words, our putative realist response begs the question against Matheson. Again, it seems that the best the realist has to go on is some kind of appeal to success to back up the claim that we are better at referring than our predecessors. The argument might go something like this:

1. If a mature scientific theory is genuinely successful, *we know* that its working posits refer and its presuppositional posits do not.

2. Most current mature scientific theories are genuinely successful.

3. Therefore, we know that the working posits of current mature scientific theories refer and its presuppositional posits do not.

Because it attempts to secure knowledge of which parts of our theories refer and which do not we can call this the 'knowledge of reference' argument. The first premise of this argument captures Kitcher's claim that "if Laudan's story has an antirealist moral, it is that the presuppositional posits of contemporary science may not exist" (Kitcher 1993, p.149). The second premise is uncontroversial so the conclusion is acceptable as long as the first premise is acceptable. Is there any reason to think that it is not?

Unfortunately, there is a way of arguing that the success of a theory is not good reason to believe that we know that the working posits of the theory refer and the presuppositional posits do not. If we claim that we know that the working posits of current scientific theories refer and their presuppositional posits do not, then it follows that many past scientists were wrong about which of their theoretical terms referred and which did not. The problem for Kitcher is that this claim allows us to reproduce the 'knowledge of reference' argument as the first three premises of a *new* pessimistic induction:

**The pessimistic induction – version 6**

1. If a mature scientific theory is genuinely successful, *we know* that its working posits refer and its presuppositional posits do not.

2. Most current mature scientific theories are genuinely successful.

3. *We know* that the working posits of current mature scientific theories refer and its presuppositional posits do not.

4. If *we know* that the working posits of most current mature scientific theories refer and its presuppositional posits do not, *then many past scientists did not know* that the working posits of most past scientific theories (sometimes) referred (i.e. when the reference of these terms is fixed causally rather than descriptively) and its presuppositional posits did not.

5. A few of these past mature scientific theories were genuinely successful.

6. If a mature scientific theory is genuinely successful, *we do not know* that its working posits (sometimes) refer and its presuppositional posits do not.

If this reformulation of Laudan's pessimistic induction is correct, then, by assuming that the success of a theory is warrant for concluding that we know that its working posits refer and that its presuppositional posits do not (statement 1), we will be able to conclude that no such warrant is available (statement 6).

The crucial thing about this argument against Kitcher is that although it employs the same basic strategy as Laudan's pessimistic induction it is not quite the same. Laudan questions the intuitive connection between success and reference by presenting successful theories whose central terms do not refer. In contrast, the argument presented above questions the intuitive connection between success and our ability to say which of our terms genuinely refer and which do not. As Matheson (1996) says, "although we can say of ourselves that we are talking about something, we can't say that we know what we are talking about" (Matheson 1996, p.473). What we have then is *another* pessimistic induction based on the fact that many past theorists have produced successful theories without knowing which of their terms referred and which did not. But is this really a "fact" about the history of science? Isn't the claim - that many past theorists have produced successful theories without knowing which of their terms referred and which did not - radically undermotivated?

As with the original formulation of Laudan's argument, there are several ways in which the realist might respond to the new pessimistic induction. Of course, given that their results are already built into the argument, the 'wrong-sort-of-success' and 'maturity' strategies are no longer available. This leaves the realist with two main ways of responding to the new pessimistic induction. The first response would be to challenge the crucial fourth premise by claiming that the inference from present knowledge (of reference) to past ignorance (of reference) does not follow. The second response would be to allow that many past scientists did not know when they were referring and when they were not, but claim that the number of such scientists who developed successful theories is insufficient to generate the pessimistic conclusion. Unfortunately, as I will now argue, there is good reason to think that neither of these responses is available to Kitcher.

The first response to the new pessimistic induction is to challenge the inference from present knowledge to past ignorance. The problem with Kitcher's account is that, in order to know which of our term-tokens refer and which do not, we have to be able to do two things. Firstly, we must be able to distinguish between the working posits of a theory and its presuppositional posits. Secondly, we must recognise that term-tokens genuinely refer only when their reference is fixed causally and not by faulty theoretical descriptions containing references to non-referring presuppositional posits. These two conditions are individually necessary and jointly sufficient for knowledge of which of our term-tokens refer and which do not. The difficulty for Kitcher is that his treatment of Laudan's most problematic examples seems to commit him to the claim that the scientists in such cases satisfied neither of the conditions for such knowledge of reference. In other words, Kitcher's account of reference supports rather than challenges the inference from present knowledge to past ignorance.

So, it seems that the first response to the new pessimistic induction is not available to Kitcher, because the scientists in Laudan's most problematic counterexamples do not satisfy the two conditions necessary for knowledge of reference. However, as suggested above, Kitcher might be able to grant premise 4 and instead question the inference from premise 5 to the conclusion. This would involve questioning the pessimistic conclusion on the grounds that there are not enough cases of ignorant-yet-successful scientists to act as an inductive foundation. Prima facie, this way of responding to the new pessimistic induction looks promising for Kitcher, because his account only requires him to say that there are a few cases at most where a scientist did not know which of the theoretical terms of a successful theory referred and which did not. But is this right? After all, in the only case Kitcher considers in any detail, his analysis requires us to say that Maxwell was wrong. What is to say that a detailed examination of other cases would not require Kitcher to say exactly the same thing about other scientists?

Of course, such questions can only really be settled by a detailed examination of the historical record. However, it should be noted that the shift in focus from questions concerning the connection between reference and success (pessimistic induction – version 1) to questions concerning the connection between *knowledge of* reference and success (pessimistic induction – version 6) has a rather unexpected consequence. Laudan's list of counterexamples was produced in an attempt to question the intuitive connection between reference and success. However, given that we are interested in cases where a theory was successful, yet the scientists who endorsed it were ignorant of which of its theoretical terms referred and which did not, we need no longer restrict our search for counterexamples to those on Laudan's list. Indeed, we are now free to look at *any* empirically successful theory, to see if there is a case to be made for the claim that the scientists who developed and endorsed it did not know which of its theoretical terms genuinely referred and which did not. Who is to say how such investigations might turn out? Is it obvious that they would support an optimistic induction from success to our knowledge of which of our terms refer and which do not?

This last point reveals something interesting about realist attempts to respond to Laudan's pessimistic induction by reconstructing his counterexamples in order to re-establish the connection between reference and success. If a consequence of such reconstructions is that a large number of past scientists did not know what they were talking about, even though their theories were successful, then they will be open to another pessimistic induction concerning *our* present ability to know which of our terms refer and which do not. If I am right about this then the only way for the realist to show that the history of science can be used to support our present knowledge of reference is to avoid large-scale reconstructions of past theoretical practice. To conclude my discussion I shall now discuss an account that at least promises such virtues.

## 8. Psillos's "divide et impera" strategy

It seems then that Kitcher's strategy of drawing a post hoc distinction between working and presuppositional posits is a lot more problematic than we might have expected. If there is a lesson to be learned from the last section then it might be that, when responding to Laudan, realists should try as much as possible to stick to distinctions that are mirrored in past scientific practice. One realist who appears to follow (or more likely anticipates) this lesson is Stathis Psillos.

Psillos (1999) follows Kitcher (1993) in thinking that "the best way to defend realism is to use the generation of stable and invariant elements in our evolving scientific image to support the view that these elements represent our best bet for what theoretical mechanisms and laws there are" (Psillos 1999, p.109). Like Kitcher, Psillos thinks that Laudan's holistic conception of theory confirmation is mistaken because it fails to allow for the possibility that some theoretical components may be responsible for the empirical success of a theory but not others:

> When does a theoretical constituent $H$ indispensably contribute to the generation of, say, a successful prediction? Suppose that $H$ together with another set of hypotheses $H'$ (and some auxiliaries A) entail prediction P. $H$ indispensably contributes to the generation of $P$ if $H'$ and A alone cannot yield $P$ and no other available hypothesis $H^*$ which is consistent with $H'$ and A can replace $H$ without loss in the relevant derivation of $P$. (Psillos 1999, p.110)

So, for Psillos, a theoretical constituent is responsible for the success of a theory only if it makes an *essential contribution* to that success. On this view, any other components of theories are 'idle' because they are "impotent to make any difference to the theory's stake for empirical success" (Psillos 1999, p.110). This is what Psillos calls the *divide et impera* move.

Is Psillos's essential vs. idle distinction the same as Kitcher's working vs. presuppositional distinction and therefore open to the same sort of objections? Psillos insists that it is not. However, he does accept that it is open to the same charge of being an ad hoc reconstruction of past theories:

> A central objection to my line thus far is the following: with the benefit of hindsight, one can rather easily work it out so that the theoretical constituents that supposedly contributed to the success of past theories turn out to be those which were, as it happens, retained in subsequent theories. So, the realists face the charge that they are bound to first identify the past constituents which have been retained and then proclaim that it was those (and only those) which contributed to the empirical success and which enjoyed evidential support (Psillos 1999, p.112)

So, unlike Kitcher, Psillos acknowledges the burden of responding to the charge that he has simply read a progressive account of theory change back into the historical record. His answer to this problem is surprising:

> In response to this objection, it should be pointed out that eminent scientists do the required identification all the time. It is not that realists come, as it were, from the future to identify the theoretical constituents of past theories that were responsible for their success. Scientists themselves tend to identify the constituents which they think were responsible for the success of their theories, and this is reflected in their attitude toward their own theories. (Psillos 1999, p.112)

If Psillos is right then it is scientists themselves who make the required distinction between the theoretical constituents that make essential contributions to the success of their own theories and the idle theoretical constituents that do not. On this view, there is no need for the realist, pace Kitcher, to impose an ad hoc distinction on the history of science in order to preserve the connection between reference and success. Scientists have already done this job for us.

The importance of Psillos's historical claim that scientists make the required distinction between essential and idle theoretical constituents of theories is that it allows him to avoid some of the problems that arise from Kitcher's ad hoc reconstruction of past theoretical practice:

> If this view is right, then not only is the *divide et impera* move not ad hoc, but it actually gains independent plausibility from the way in which scientists treat their theories, and from the way they differentiate their commitments to their several constituent theoretical claims. (Psillos 1999, p.113)

So, if the divide et impera move is acceptable, Psillos has a powerful *historical* argument in favour of the reference preserving abilities of scientists that does not rely on an ad hoc historical reconstruction. However, Psillos realizes that his historical promissory note needs cashing out with some proper examples. So, in order to defend the claim that scientists make the distinction between essential and idle theoretical constituents in a way that supports realism, he attempts to show that this is even the case in two of Laudan's most persuasive examples of successful theories whose central terms were non-referential: the caloric theory of heat and the optical ether theories of the nineteenth century.

Psillos's major conclusions about the caloric theory of heat can be summarized as follows:

1. The laws which scientists considered well supported by the available evidence and the background assumptions they used in their theoretical derivation were *independent* of the hypothesis that the cause of heat was a material substance: no relevant assumption was essentially used in the derivation-prediction of these laws.

2. The parts of caloric theory which scientists believed in were well supported by the evidence and were retained in subsequent theories of heat, whereas the hypotheses that were abandoned were those which were ill-supported by the evidence.

3. The caloric representation of the cause of heat as a material fluid was not as central, unquestioned and supported as, for instance, Laudan (1984a:.113) has claimed. Caloric was not a putative entity to which the most eminent scientists had committed themselves as the real causal agent of heat phenomena.

(Psillos 1999, p.113)

The first claim here is essentially Kitcher's argument for presuppositional posits in a slightly altered form, namely that suspect theoretical posits were a dispensable part of caloric theory and therefore did not contribute to its success. However, Psillos's real advance over Kitcher is in the third claim that scientists themselves recognized the dispensability of caloric at the time. In contrast to Kitcher, Psillos argues not only for the *logical* dispensability of suspect theoretical posits but also for their *historical/psychological* dispensability as well.

Psillos's second case study concerns the role of the ether in nineteenth century optical theories. The major conclusions of this second case study can be summarized as follows:

1. The investigation of the possible constitution of the carrier of light-waves was aided by the construction of models (e.g. Green's elastic-solid model of the ether), where this model construction was based on perceived analogies between the carrier of light-waves (e.g. its ability to sustain transversal waves) and other physical systems (e.g. elastic solids). It was mostly these models that were abandoned later on.

2. The most general theory – in terms of Lagrangian dynamics and the satisfaction of the principle of the conservation of energy – which was the backbone of the research programme around the dynamical behaviour of the carrier of light-waves has been retained in the subsequent framework of electromagnetism.

3. The parts of 'luminiferous ether' theories which were taken by scientists to be well-supported by the evidence and to contribute to well-founded explanations of the phenomena were retained in subsequent theories. What became paradigmatically abandoned was a series of models which were used as heuristic devices for the possible constitution of the ether.

(ibid. p.140)

These three claims correspond to the claims made in the caloric theory case study. The first claim concerns the dispensability of various models of the ether as the carrier of light-waves. The second claim concerns the retention of well-supported theories and the subsequent rejection of suspect models of the ether based on the analogy with elastic solids. Again, Psillos's advance over Kitcher is in the third claim that those parts of ether theories that have been abandoned in the subsequent framework of electromagnetism were not thought to be accurate descriptions of the ether by the scientists who developed them.

So, in both of his case studies Psillos thinks he can show that scientists are able to recognise which of their theoretical posits refer and which do not. However, it should be noted that there is a crucial difference between the cases. For although Psillos is happy to say that 'caloric' did not refer he seems reluctant to make a similar claim about 'ether'. Thus, in order to preserve the claim that the central theoretical terms of successful scientific theories genuinely refer, Psillos develops a causal-descriptive theory of reference according to which we can view terms-tokens of 'ether' as referring to the electromagnetic field. Like Hardin and Rosenberg (1982), Psillos argues that a causal theory of reference is attractive because it allows us to maintain continuity of reference through periods of conceptual upheaval. However, Psillos also argues that a purely causal approach to such issues is inadequate for at least two reasons. Firstly, the causal theory of reference is unable to explain how we are able to refer to natural kinds. Here Psillos argues that our ability to refer to natural kinds can only be explained by making use of theoretical descriptions whereby "the burden of reference is borne by the *kind-constitutive properties* attributed to the posited kind, substance or object" (Psillos 1999, p.287). Secondly, the causal theory makes our ability to refer completely trivial because "it is not clear what could possibly show that the relevant theoretical term does not refer. If the reference of the term is fixed purely existentially, then insofar as there *is* a causal agent behind the relevant phenomena, the term is bound to end up referring to it" (ibid. p.290). Thus, in order to explain referential failure as well as referential success, Psillos concludes "the burden of reference of theoretical terms lies with *some* descriptions which specify the kind-constitutive properties by virtue of which the referent, if it exists, plays its causal role" (ibid. p.291).

Psillos's argument for the necessity of a causal-descriptive theory of reference looks very much like Kitcher's, but it is important to realise that their respective ideas of what such a theory should look like and what it can hope to achieve are not the same. Kitcher's approach to apparently non-referring central theoretical terms is based on showing that: 1) such terms were presuppositional rather than central, and; 2) such non-referring presuppositional posits only affect our ability to refer when reference is fixed descriptively. In contrast, Psillos argues that: 1) such posits were acknowledged to be idle by scientists themselves, or: 2) such posits were central but they referred because "the term which is employed to denote the posited entity is associated with a *core causal description* of the properties by virtue of which it plays its causal role *vis-à-vis* the set of phenomena" (ibid. p.295). Thus, for Psillos, but not Kitcher, it is possible for a theoretical term to get its reference fixed descriptively even though it is associated with a number of false theoretical descriptions. What is important for Psillos is that the core causal descriptions that fix the reference of 'ether' are the same as those that fix the reference of 'electromagnetic field.' This is why Psillos can say, pace Kitcher, that the ether referred to the electromagnetic field.

What are we to say about Psillos's attempt to preserve the connection between reference and success? Does it fair any better than Kitcher's account as a response to the pessimistic induction? The first thing to say about Psillos's approach is that by focussing on the *actual* attitudes of scientists it promises to avoid the charge that faces Kitcher's attempt to distinguish between working and presuppositional posits, namely the charge of being an ad hoc rational reconstruction of past theoretical practice. Unlike Kitcher, Psillos is keen to show that his distinction between essential and idle theoretical posits is historically realised, not simply a philosopher's fiction. This focus on the actual attitudes of scientists also allows Psillos to defend his causal-descriptive theory of reference from the charge that the identification of 'core causal descriptions' is similarly ad hoc. If Psillos is right then the core causal descriptions of terms like 'ether' are apparent in the way past scientists were able to distinguish between the theoretical descriptions that were well-supported and those that were suspicious or merely heuristic devices for investigating other phenomena. In other words, past scientists not only make the requisite distinction between essential and idle theoretical posits, they also make the requisite distinction between essential and idle (or perhaps heuristic) theoretical descriptions.

So, according to Psillos, if one looks hard enough a realist account of science as converging on the truth is there in the pages of history. On this view, the reconstructive (ad hoc) tendencies of old-school realists like Kitcher are mistaken attempts to do something that cannot and, indeed, need not be done. An important consequence of this claim is that Psillos will not only be able to block the original pessimistic induction, he will also be able block the new version of the pessimistic induction discussed in the last section by denying the crucial fourth premise that relies on the inference from present knowledge to past ignorance. If the conclusions of Psillos's case studies are acceptable then we will have no grounds for this inference and, *a fortiori*, no grounds for the pessimistic conclusion that we don't know which of our terms genuinely refer and which do not. Does this mean then that Psillos has finally refuted the pessimistic induction?

## 9. Response to Psillos

Whether or not one is convinced by his case studies one must at least congratulate Psillos for making it clear that the dispute between opponents and supporters of the pessimistic induction must be fought on *historical* grounds. Like Giere (1988), Psillos accepts that realism must be defended on descriptive as well as normative grounds.[4] In contrast to most of his realist predecessors (i.e. Hardin and Rosenberg, Boyd, Kitcher, etc.) Psillos recognises that simply creating ad hoc rational reconstructions that whiggishly support a realist reading of science cannot appease the worries raised by Laudan. Indeed, Psillos appears to acknowledge that such tactics will result in something like the argument discussed in the previous section according to which we can generate a new pessimistic induction concerning our (likely) ignorance of which of our terms refer and which do not.

However, the sensitivity of Psillos's account to the historical plausibility of realism is also its major weakness. By putting so much argumentative weight on the results of his historical case studies Psillos presents the antirealist with an obvious course of action, namely to question his substantive historical claims. Here the antirealist could make use of a variety of strategies developed by historians, sociologists and philosophers of science. Firstly, and most obviously, we could re-examine the historical evidence to see whether or not it unequivocally supports Psillos's realist reading. Secondly, we could question Psillos's reliance on first-person reports as transparent descriptions of scientists' beliefs. Thirdly, we could question his narrow focus on theory as the final arbiter in what counts as an essential theoretical posit or description. For example, we might try to show that although caloric was an idle constituent of theory it was not an idle constituent of practice.

By suggesting these ways of responding to Psillos's historical defence of realism I do not mean to claim that they would necessarily be successful at undermining the major claims of his caloric and ether case studies. They are put forward simply to demonstrate that matters are not quite as straightforward as Psillos would have us believe. However, given the amount of evidence Psillos presents in favour of a realist interpretation of his case studies, can't we at least say that the burden of proof is now on the antirealist to show that such interpretations cannot be sustained? Perhaps, but it is important to realize that in questioning Psillos's historical argument for realism we need not restrict ourselves to his case studies nor even to the examples on Laudan's list. As I suggested at the end of the previous section, the problem for the realist is that once we move to the question concerning the inference from success to *knowledge of* reference it is entirely legitimate for the antirealist to re-examine the history of successful theories. Just because the central theoretical terms of a successful theory refer (by our standards) it does not follow that the scientists who endorsed and developed the theory were aware of this 'fact'.

Is it possible to identify cases where scientists were not aware of which of their theoretical posits referred yet they were able to develop successful theories? In her recent book, *Social Empiricism*, Miriam Solomon (2001) argues for a position she calls 'whig realism,' the main claim of which is that a realist account of the history of science is *only* possible with the benefit of hindsight. Thus, according to Solomon, all we can say about past successful theories is that there is *some truth* in them where judgments of truth and success are admittedly whiggish evaluations based on current standards. Now it could be argued that with friends like Solomon realism is hardly in need of enemies. However, it is not my intention here to assess the viability of Solomon's account but rather to draw attention to some of the examples she uses to support whig realism. Many of these examples demonstrate just how difficult it is to square what *we* think of past successful theories with the *actual* judgments of past scientists. To take just one of example:

> "Phlogiston" will turn out to have reference at least sometimes because "dephlogisticated air" was successfully – *albeit unwittingly* – used in a particular set of experimental contexts to refer to oxygen (Kitcher 1993, p.100). That is, "phlogiston" sometimes refers even though there is no such thing as Priestley's phlogiston, and even though no phlogiston theorist at the time could say how the term refers. Phlogiston theorists were actually often successfully referring to oxygen when they used the phrase "dephlogisticated air," although they did not know it. (Solomon 2001, p.37)

Here then it seems we have a case of a successful past theory whose central theoretical terms referred even though the scientists who developed and used such terms did not

---

[4] Giere supports his realist account of science with laboratory studies that attempt to show that realism is a necessary presupposition of scientific practice. See Giere (1988) chapter 5 - 'Realism in the Laboratory' for further details of this argument.

know how or even that they referred. In other words, we seem to have a counterexample to Psillos's claim that past scientists were able to distinguish between well supported as opposed to theoretically suspect parts of their theories.

Prima facie, there are two ways in which Psillos can respond to the phlogiston example. He could either follow Kitcher (1993) by arguing that term-tokens of "dephlogisticated air" genuinely refer when their reference is fixed causally or, alternatively, he could argue that such term-tokens refer because their reference is fixed by a core causal description. The problem for Psillos is that, on his own terms, neither of these responses will result in a view of past theoretical practice that supports his historical argument for realism. The Kitcherian response won't do because although it allows us to say that "phlogiston" refers it also requires us to say that it did so *albeit unwittingly*. Similarly, the core causal description response that Psillos favours in the ether case study won't do either because although it might be possible to say that some of the theoretical descriptions associated with "phlogiston" and "dephlogisticated air" have been retained in subsequent theories, we will also have to say that scientists at the time did not identify these descriptions as better confirmed than those we have subsequently rejected.

So, it seems that in order to save his account of the connection between success and knowledge of reference Psillos will have to say that terms like "phlogiston" and "dephlogisticated air" fail to refer. Indeed, this is precisely what we find:

> Joseph Priestley and other advocates of the phlogiston theory were causally connected to oxygen – they breathed it and it was causally involved in the experiments they made to investigate combustion. But none of the properties of oxygen were the causal origin of the information they had associated with phlogiston. And nothing in nature could possibly be the causal origin of such information. What it is correct to say is that 'phlogiston' refers to nothing. But in order to say this, we need to say that there is nothing in nature which possesses the properties that phlogiston was supposed to possess. (Psillos 1999, p.291)

Presumably the reason that Psillos is motivated to say that 'phlogiston' does not refer is that he realises what the consequences would be for his historical argument for realism if he were to claim that it did. But there's the rub because by denying that 'phlogiston' refers Psillos now owes us an account of how we can have a successful scientific theory whose central theoretical term does not refer. Now we know that Psillos cannot claim that 'phlogiston' referred after all because this will also require us to say that past scientists referred even though they did not know how or that they were doing so. So, the only option remaining is for Psillos to deny that phlogiston theory was not successful. Is this a viable claim?

Unfortunately, in the debate between Laudan and his critics even the majority of realists have largely accepted that the phlogiston theory is an example of a successful scientific theory whose central terms fail to refer. This is precisely why there are so many attempts in the literature to show how 'phlogiston' and 'dephlogisticated air' may be seen as referring terms. Of course, Psillos is free to define success so that it turns out that the phlogiston theory was not successful, but how could we avoid the suspicion that this is simply an ad hoc manoeuvre to eliminate problematic examples? It seems then that Psillos cannot win. If he says that 'phlogiston' *does refer* he must face the charge that his identification of a core causal description is an ad hoc rational reconstruction. If he says that 'phlogiston' *does not refer* he must face the charge that he has gerrymandered his definition of success in order to maintain the historical connection between success and knowledge of reference. Either way he has failed in his aim to refute the pessimistic induction.

Psillos might object to this argument on the grounds that, even if correct, it is just one counterexample. Accordingly, he could say that the phlogiston example is the exception to the rule that success implies knowledge of which of terms refer and which do not. Unfortunately, there are two problems with this response. Firstly, it is not clear that the phlogiston example is the only example of a successful theory that fails to fit Psillos's argument. Secondly, even if it were the only counterexample available, it is not clear that the realist can allow exceptions to the connection between success and (knowledge of) reference. Original formulations of the explanationist defence were based on the idea that the connection between success and reference should allow no exceptions (hence the claim that it would be a *miracle* if the central theoretical terms of successful theories did not refer). On occasion Psillos hints that realists should no longer commit themselves to such a strict connection but he never provides any justification for this claim.

## 10. Conclusion

In this chapter we have seen a number of ways in which realists have attempted to undermine Laudan's pessimistic induction. I have argued that although such strategies ultimately succeed in undermining the original formulation of Laudan's argument they result in a new pessimistic induction concerning our ability to know which of our theoretical terms genuinely refer and which do not. If I am right then it is not enough for the realists like Kitcher to provide rational reconstructions of past theoretical practice. In order to defeat the new pessimistic induction the realist needs to show that the success of a theory implies *knowledge of* reference. I have shown that although Psillos (1999) makes some progress in this direction he ultimately falls back on ad hoc rational reconstructions when it comes to Laudan's most problematic example. However, in the next chapter I will consider a response to the pessimistic induction that suggests that none of the evidence Laudan presents, *no matter how it is interpreted*, can support his pessimistic conclusion.

# Chapter 4

# A Defence of Laudan's Pessimistic Induction: Part 2

## 1. Introduction

What if we could accept Laudan's counterexamples and yet still maintain that success is a reliable indication that a theory is true? This is the alternative offered by Peter Lewis (2001) who argues that the pessimistic induction commits a well-known probabilistic fallacy. According to Lewis, the main problem with Laudan's argument is that the pessimistic conclusion does not follow from his list of counterexamples. Based on an analogy with diagnostic testing for disease Lewis suggests that success can be a reliable test for truth even though some successful theories are false and some unsuccessful theories are true. In this way we are offered a way out of the impasse generated by attempts to put a realist spin on Laudan's counterexamples. If Lewis is right then there is no need for such reconstructions because these so-called counterexamples are, and always were, perfectly compatible with scientific realism. The aim of this chapter is to decide whether or not this new way of responding to Laudan's argument is a viable alternative to the standard realist response.

In the next section, I outline Lewis's argument showing how it is based on an analogous argument for the reliability of diagnostic testing for disease. In section 3, I suggest that the best way to assess this argument is to take Lewis's analogy seriously by comparing an example of diagnostic testing for disease with the success/truth case. Here I argue that an example suitable for this purpose must meet four conditions. In section 4, I put forward diagnostic testing for Alzheimer's disease as an example that meets these conditions. In section 5, I argue that there is a major disanalogy between this example and the success/truth case that shows why Lewis's argument cannot work as a response to Laudan's pessimistic induction. This is the problem of independent testing. In section 6, I introduce the direct and indirect methods of determining truth as two possible ways in which we might attempt to rescue Lewis's argument from this problem. However, I argue that neither of these methods is suitable for achieving this aim. Finally, in section 7, I conclude by suggesting that the problems that affect Lewis's argument may indicate the need for a more radical approach to the issues surrounding Laudan's pessimistic induction.

## 2. The false positives argument

Lewis (2001) reconstructs Laudan's pessimistic induction as a reductio that takes the following form:

(1) Assume that the success of a theory is a reliable test for its truth.

(2) Most current scientific theories are successful.

(3) So most current scientific theories are true.

(4) Then most past scientific theories are false, since they differ from current theories in significant ways.

(5) Many of these false past theories were successful.

(6) So the success of a theory is not a reliable test for its truth.

(Lewis 2001, p.373)

Unlike his predecessors who have tended to focus on premises (4) and (5), Lewis suggests that the real problem with Laudan's argument is to be found elsewhere:

> The fallacious move is between premise (5) and the conclusion; that many false past theories were successful does not warrant the assertion that success is not a reliable test for truth. (Lewis 2001, p.374)

This claim is likely to strike many as counterintuitive. How can success test for truth if we allow that strictly false theories can be successful?

Lewis begins his justification for this claim by making some observations about the notion of reliability:

> Consider the standard characterization of the reliability of a test in terms of the rates of false positives and false negatives. For example, for a diagnostic test for some disease, the false positive rate is the probability that the test gives a positive result given that the patient does not have the disease, and the false negative rate is the probability that the test gives a negative result given that the patient does have the disease. Since these are two ways in which the test can go wrong, a reliable test is one in which the false positive rate and false negative rate are both sufficiently small, where what counts as sufficiently small is determined by the context. (Lewis 2001, p.375)

So, in asserting the reliability of a diagnostic test for a particular disease we are not committed to saying that there are absolutely no occasions when the test lets us down but rather that the number of times it does so is sufficiently small for us to rely on the test as a diagnostic procedure. In other words, the reliability of any particular test is compatible with there being certain occasions when the test lets us down.

There is nothing particularly controversial about Lewis's characterization of how the reliability of diagnostic medical tests is established. However, what certainly is controversial is his further claim that it is precisely the same notion of reliability at work in the explanationist defence of scientific realism:

> Essentially, the realist maintains that success can be used as a test of the truth of theories, since we can directly observe the success of a theory, but not its truth. For such a test, a false positive would be a case in which a theory is false but successful, and a false negative would be a case in which a theory is true but unsuccessful. In affirming that success provides a reliable test for truth, the realist is claiming that the rates of false positives and false negatives are low. If this is indeed the case, then if most current scientific theories are successful, it follows deductively that most current theories are true. (Lewis 2001, p.375)

Lewis is not alone in stressing that the realist should allow for a certain number of cases where false theories are successful and true theories are unsuccessful. As Psillos (1999, p.80) says, "the realist argument should acknowledge the existence of failures. Their actuality does not impair scientific methodology." However, as we will now see, Lewis differs from Psillos and most other realists in that he thinks this can be shown without the need for historical reconstructions of Laudan's putative counterexamples.

Having stipulated that the notion of reliability behind the explanationist defence is the same as the operative notion of reliability that supports diagnostic tests for disease, Lewis is able to offer us a fresh interpretation of Laudan's critique:

> Laudan argues that evidence from the history of science should undermine our confidence in the reliability of success as a test for truth. Presumably this means that the historical record gives us reason to believe that either the false positive rate or the false negative test for the test is large. (Lewis 2001, p.375)

So, according to Lewis, when Laudan tells us that there are many examples in the history of science of false but successful and true but unsuccessful scientific theories he is actually claiming that the false positive rate (theory false but test for success positive) and the false negative rate (theory true but test for success negative) are high. However, Laudan must do more than just show that there are *some* occasions where the test for truth fails. In order to show that success is an unreliable test for truth, Lewis demands that Laudan (or for that matter any opponent of the explanationist defence) show that there are a *sufficient number* of cases where the test fails.

It seems obvious that this is not the way Laudan (1996a) would choose to frame his dispute with the realist. As suggested earlier, Laudan primarily aims to show that truth is neither a necessary nor sufficient condition of the success of science.[1] Thus, for Laudan at least, the issue of how many counterexamples he can present would seem to be irrelevant. However, as Lewis points out, there are places where Laudan (1996a) provides a rough indication of how many counterexamples we might expect to find in the history of science:

> The strongest statement of his historical evidence that Laudan makes is the following: "I daresay for every highly successful theory in the past of science which we now believe to be a genuinely referring theory, one could find half a dozen once successful theories which we now regard as substantially non-referring" (Laudan 1981, p.35). His claim is that if we judge the truth of theories according to whether we now believe their central terms to refer, then among past successful theories the false ones outnumber the true ones six to one. (Lewis 2001, pp.375-6)

So Lewis attributes to Laudan the claim that for every true positive (a theory that is both true and successful) in science there are six false positives (theories that are false but successful). Although this is a rather strange way of interpreting Laudan's views, Lewis is surely correct in thinking that Laudan takes the 1:6 ratio to show that the explanationist defence of scientific realism is unworkable.

Let us grant then that it makes sense to attribute to Laudan the claim that for every true positive there are six false positives. Lewis wants to argue that this ratio provides insufficient evidence from which we can conclude that success is an unreliable test for the truth of scientific theories. The basis of this claim again comes from the connection he makes between the explanationist defence and diagnostic testing for disease. Here Lewis asks us to consider a highly reliable test for a particularly rare disease. The test is reliable because the false positive and false negative rates are both

[1] It should be noted that there is a certain asymmetry in Laudan's discussion. Whereas lots of good examples of successful but non-referring theories are introduced, he only proposes a few examples of a genuinely referring but unsuccessful theory, e.g. chemical atomic theory in the 18th century, Proutian atomic theory, and continental drift theory in the 1930s.

only one in a hundred. The disease is rare because only one person in ten thousand actually has it. So if we test a particular person at random the chances that the person does not have the disease and tests positive for it are one in a hundred, and the chances that they have the disease and test positive for it are one in ten thousand. So for every true positive test result we should expect one hundred false positive results – a ratio of 1:100. The *false positives paradox* is that this ratio of true positives to false positives is perfectly compatible with the reliability of the test in question.

If the explanationist defence relies on the same notion of reliability as that employed for diagnostic testing then it is very easy for Lewis to show that a ratio of one true positive to six false positives is perfectly compatible with the reliability of success as a test for truth. He does this by making two assumptions that mirror those made in the case of the rare disease:

**Assumption 1.** True theories are relatively rare, e.g. 1 in 25 theories are true.

**Assumption 2.** The test for true theories is reliable, e.g. the false positive rate and false negative rate for success as a test for truth are both relatively low, e.g. 1 in 5.

In this scenario the chances of a theory being true and successful will be 1/25 x 4/5 = 4/125, and the chances that it is false but successful will be 24/25 x 1/5 = 24/125. In other words, the ratio of true positives to false positives will be precisely the 1:6 ratio put forward by Laudan. For Lewis, this shows that Laudan's list of false but successful scientific theories is not as harmful to realism as one might think:

> The realist can interpret Laudan's historical cases, not as evidence against the reliability of success as a test for truth, but merely as evidence of the scarcity of true theories in the past. (Lewis 2001, p.377)

Laudan allegedly falls pray to the false positives paradox by overlooking the fact that a reliable test for truth can produce a high ratio of false positives to true positives as long as true theories are rare. But, of course, this last point is precisely what the convergent realist wants to claim. Thus, according to Lewis, Laudan's list of counterexamples not only fails to show that success is an unreliable test for truth; it in fact provides indirect evidence for convergent realism.

The ingenious part of Lewis's response to the pessimistic induction is that it shows how a realist can accept all of the historical evidence Laudan (1996a) presents but still hold that success is a reliable test for truth:

> I do not wish to suggest that Laudan's project of subjecting the claims of the convergent realist to historical scrutiny is ill-motivated; my claim is only that the historical evidence he provides is not adequate to the task. (Lewis 2001, p.379)

Rather than questioning Laudan's *interpretation* of the historical evidence, Lewis instead questions the *adequacy* of that evidence for the conclusion it is supposed to support. His reason for doing this again results from the analogy he draws between the success/truth case and the diagnostic test/disease case:

> In the case of a test for a disease, the most straightforward way of undermining the claim that the test is reliable would be to take a random sample of people who are free of the disease, and show that a significant proportion of them test positive. Analogously, the most straightforward way to undermine the claim that success is a reliable test for truth would be to take a random sample of theories which are known to be false, and show that a significant proportion of them are nevertheless successful. (Lewis 2001, p.379)

Thus, if Lewis is right, Laudan has got things the wrong way round. Simply showing that there are false but successful theories in the history of science cannot show that success is not reliable test for truth. To do this Laudan needs to show that a *significant proportion* of false theories were successful.

## 3. Tests for truth vs. Tests for disease

If the probabilistic reading of the pessimistic induction is right then it seems that the realist can rest easy. Laudan simply does not have enough evidence to support his pessimistic conclusion. However, before rushing off to the history books to look for fresh counterexamples it would surely be wise to check the validity of the false positives argument as applied to the reliability of success as a test for truth. One obvious way to do this is to examine the analogy Lewis draws between success as a test for truth and diagnostic testing for disease. Initially, this would seem to present us with two questions:

1. Does success test for truth in the same way that diagnostic tests detect the presence of disease?
2. Is the truth of scientific theories similar to the presence of a disease in a human body?

Unfortunately, Lewis does not provide us with enough information to properly answer these questions because his discussion relies on vague references to "a diagnostic test for some disease" (Lewis 2001, p.374) without giving any specific examples. If the analogy were merely illustrative then this vagueness would be understandable but given that it bears the justificatory weight of Lewis's argument it does not seem unreasonable to ask for details. If success is like a diagnostic test, what kind of diagnostic test is it like? If truth is like a disease, what kind of disease is it like?

In order to assess the acceptability of the analogy between the success/truth case and the test/disease case we need to know more about the kind of example Lewis takes to representative of the latter. Given that Lewis does not provide us with details of the specific tests or diseases he has in mind it seems that we must rely on the following characterization of the success/truth case:

> Essentially, the realist maintains that success can be used as a test of the truth of theories, since we can directly observe the success of a theory, but not its truth. (Lewis 2001, p.375)

The value of this formulation is that it gives us crucial information about the features Lewis thinks the test/disease case must have in order for it to be analogous to the success/truth case. Thus, according to Lewis, this analogy holds only if:

**Condition 1.** The diagnostic test in question tests for symptoms that are directly observable.

**Condition 2.** The disease in question is unobservable.

These two conditions place obvious constraints on the sort of example we might use to assess the acceptability of Lewis's analogy. For although there are many examples of the test/disease case that will satisfy both of these conditions, it is important to note that there are also many examples that will not.[2]

If a particular case of diagnostic testing for disease is to be a suitable way of cashing out Lewis's analogy then it seems obvious that it must support the false positives argument outlined in section 2. However, the fact that a particular set of diagnostic tests identifies observable symptoms (Condition 2) in order to detect the presence of a particular unobservable disease (Condition 1) tells us nothing about whether or not this example will support the false positives argument. This can only be decided by knowing more about the case in question. This is clear from Lewis's characterization of the sort of features his argument requires:

> Consider again a diagnostic test for some disease, and let us stipulate that the test is highly reliable; for example, suppose that the false positive rate and the false negative rate are both one in a hundred. Further suppose that the disease in question is rare – that only one person in ten thousand has it. (Lewis 2001, p.376)

In other words, a particular case of testing for disease will only support the false positives argument if the test and disease in question satisfy the following two conditions:

**Condition 3.** The diagnostic test in question must be *reliable*.

**Condition 4.** The disease tested for must be *rare*.

Unlike conditions 1 and 2, which focus on the qualitative feature of observability, these conditions place quantitative constraints on the selection of an appropriate example for Lewis's analogy. In particular, we must be able to show that we can produce figures for reliability and rarity such that false positives will outnumber true positive results in the general population.

Combining all four of the above conditions finally puts us in a position to say what kind of example is likely to fit Lewis's analogy. This example would be a case where we use a *reliable* test to identify directly *observable* symptoms that are caused by the presence of an *rare* and *unobservable* disease, where the latter can be anything from the breakdown of normal functioning at a particular anatomic site (e.g. a cancer lesion)

---

[2] Many diseases are directly observable just as many tests for disease attempt to detect unobservable symptoms. For an excellent philosophical discussion of the various ways in which disease is classified and tested for see Stempsey (2000).

to the presence of a specific etiological agent (e.g. a virus). In the next section I will present an example of diagnostic testing for disease that has just these features.

## 4. Cashing out the analogy: Alzheimer's disease

Alzheimer's disease is a common form of dementia that impairs normal cognitive functioning. Symptoms include memory loss, difficulty in speaking, and an inability to perform everyday tasks. Although its ultimate cause is unknown, post-mortem analysis reveals that the brain tissue of all sufferers contains abnormal clumps (amyloid plaques) and tangled bundles of fibres (neurofibrillary tangles). At the present time post-mortem identification of 'clumps and tangles' is the only way to conclusively diagnose the disease. However, in order to help sufferers and their carers, doctors have devised several tests to enable a diagnosis of 'probable' Alzheimer's disease. These include medical tests of blood, urine or spinal fluid; neuropsychological tests that measure memory, problem solving, attention, counting and language; and brain scans to check for abnormalities. Using such tests doctors can correctly diagnose Alzheimer's disease 90 percent of the time.[3]

My suggestion is that Alzheimer's disease is paradigmatic of the kind of case Lewis has in mind when he says that the success/truth case is analogous to "a diagnostic test for some disease" (Lewis 2001, p.375). There are two features of the Alzheimer's case that explain why it is such a good candidate for Lewis's analogy. Firstly, most of the tests used to diagnose 'probable' Alzheimer's disease are designed to detect the appearance of directly observable symptoms, e.g. loss of cognitive abilities. So the test for probable Alzheimer's disease satisfies the first condition for the applicability of Lewis's analogy.[4] Secondly, the cause of the observable symptoms, i.e. the disease, is itself unobservable. This is the case whether we choose to define the disease causally as the (unknown) agent or mechanism behind the production of observable and unobservable symptoms or anatomically as the (known) abnormal structures found in brain tissue (clumps and tangles). So the disease also satisfies the second condition for the applicability of Lewis's analogy.

[3] My brief summary of Alzheimer's disease is largely based on the online fact sheet that can be found at http//www.alzheimers.org/pubs/adfact.html.

[4] It might be suggested that diagnostic testing for Alzheimer's disease does not meet this condition given that at least some of these tests, e.g. brain scans, are designed to detect 'unobservable' symptoms. I have chosen to ignore this possible objection for two reasons. Firstly, tests for unobservable symptoms are still

Whether or not diagnostic testing for Alzheimer's disease meets condition 3 is an empirical matter that can only be decided by assessing the relevant data concerning the reliability of testing. As noted above, using tests based on standard criteria specialized centres can correctly diagnose Alzheimer's disease about 90% of the time. However, as Blacker (1994) has shown, the specificity and sensitivity of the criteria used to diagnose Alzheimer's, NINCDS-ADRDA, are only 0.81 and 0.73 respectively.[5] Thus, the false negative rate and false positive rate for tests based on the NINCDS-ADRDA criteria are 0.19 and 0.27 respectively. However, in Lewis's initial characterization of the false positives argument he introduces an example of testing for disease where the false positive rate and false negative rate are both 0.01. The problem is that when judged by this standard of reliability diagnostic testing for Alzheimer's disease does not seem to be particularly reliable. [6] More importantly, it seems to show that this kind of diagnostic testing is not reliable enough to meet condition 3.

Unfortunately, the same line of reasoning would also allow us to conclude that success is not a reliable test for truth since the false positive and false negative rate Lewis puts forward for this case are in fact very similar to those in the Alzheimer's case:

> Suppose that 1 in 25 theories are true, and that the false positive rate and false negative rate for success as a test for truth are both 1 in 5. Such rates are consistent with success being a reasonable test for truth. (Lewis 2001, p.376)

If a false positive rate of 0.2 and false negative rate of 0.2 (1 in 5) are consistent with success being a reasonable test for truth then can we not also say that a false positive rate of 0.27 and a false negative rate of 0.19 show that tests based on NINCDS-ADRDA criteria are reasonable ways of detecting Alzheimer's disease? Lewis tells us that these figures are compatible with success being a reasonable test for truth. This shows that the standard of reliability we must use in selecting a suitable example for Lewis's analogy is the weak one given in the success/truth case rather than the much stronger one given in the medical example. For although diagnostic testing for Alzheimer's disease may be unreliable when judged by the standards of the medical community it is sufficiently

---

a peripheral part of diagnosing Alzheimer's disease. Secondly, it would be hard to find any disease that is diagnosed purely on the basis of observable symptoms.

[5] For more on the NINCDS-ADRA criteria see Jorm (1990) and Blacker (1994).

[6] This is also the conclusion of Blacker (1994).

reliable to be analogous to Lewis's characterization of success being a reliable test for truth.

We now turn to the fourth condition concerning the applicability of the false positives argument. Is Alzheimer's disease rare enough for the false positives argument to work? Prima facie, it would appear not because, as suggested above, Alzheimer's disease is the most common form of dementia in elderly people. However, the fact that Alzheimer's disease is a common form of dementia is not to say that it is not sufficiently rare for it to meet our fourth condition.[7] The only way of settling this issue is to look at how common the disease is in the general population; table 1 presents the prevalence of Alzheimer's disease in four separate age groups.

**Table 1** Age-specific prevalence rates (per 100 population) for Alzheimer's disease in Europe. From Rocca *et al.* (1991)

| Age (years) | Prevalence (per 100 population) |
|---|---|
| 35-69 | 0.2 |
| 60-69 | 0.3 |
| 70-79 | 3.2 |
| 80-89 | 10.8 |

What we want to know now is whether or not we can combine these figures with a value for the reliability of testing for AD to produce the kind of ratio between false positives and true positives required by Lewis's version of the false positives argument. Is this possible?

Combining the figures in table 1 with the false positive rate of 0.27 for tests based on the NINCDS-ADRDA criteria allows us to calculate the ratio of false positives to true positives in each particular age group. The results of this analysis are presented in table 2.

**Table 2** Age-specific ratios of false positives to true positives (per 100 population) for Alzheimer's disease in Europe. Based on false positive rate of 0.27 for NINCDS-ADRDA criteria (from Blacker 1994).

[7] Disease x can be both rare and the most common form of condition y as long as the latter is not itself a common condition.

| Age (years) | Prevalence (per 100 population) | True positive test results (per 100) | False positive test results (per 100) | Number of false positives to every true positive |
|---|---|---|---|---|
| 35-69 | 0.2 | 0.15 | 26.95 | 184.56 |
| 60-69 | 0.3 | 0.22 | 26.92 | 122.92 |
| 70-79 | 3.2 | 2.34 | 26.14 | 11.19 |
| 80-89 | 10.8 | 7.88 | 24.08 | 3.05 |

Table 2 shows that in each age group we can expect diagnostic tests based on NINCDS-ADRDA criteria to yield more false positive results than true positive results. This shows that Alzheimer's disease is rare enough to support Lewis's false positives argument. All that is required for this argument to work is for false positives to outnumber true positives, so the fact that this ratio is significantly higher in younger age groups than in older age groups is irrelevant. Alzheimer's disease meets condition 4 for the applicability of the false positives argument, *no matter which age group is selected.*

We have spent the last two sections attempting to put some flesh on the bones of Lewis's analogy. We did this by selecting an example of diagnostic testing for disease that met four conditions: observability of symptoms tested for, unobservability of disease, reliability of test, and rarity of disease. Insofar as diagnostic testing for Alzheimer's disease meets these conditions we can say that it is a suitable way of cashing out Lewis's analogy.

## 5. The problem of independent testing

We are now in a position to assess Lewis's claim that the false positives argument can be used to defend the reliability of success as a test for truth. This claim effectively amounts to showing that the success/truth case has the same kind of features as the Alzheimer's disease case. In other words we must be able to show that it meets the following four conditions:

**Condition 1'.** The truth of a scientific theory is unobservable.

**Condition 2'.** The success of a scientific theory is observable.

**Condition 3'.** The success of a scientific theory is a reliable test of its truth.

**Condition 4'.** True scientific theories are rare.

The first thing we might want to say here is that although success is surely observable, it seems strange to characterize truth as being unobservable. What can this mean? Is truth an unobservable thing or a process? If I had better senses or a bigger microscope could I see it for myself? These are difficult questions and they certainly need answering if we are to make sense of the claim that truth is an unobservable property of scientific theories. However, I want to put to one side these weighty issues concerning the nature of truth and instead focus on conditions 3' and 4'.

It may seem obvious that the success/truth case satisfies conditions 3' and 4'. After all when selecting the Alzheimer's example we explicitly used Lewis's characterization of the success/truth case as a benchmark in deciding when a test and a disease were respectively reliable and rare enough to allow for the application of the false positives argument. We did this by comparing known figures for the reliability of diagnostic testing for Alzheimer's disease and figures for the prevalence of this disease with the figures given by Lewis for the reliability of success as a test for truth and the prevalence of true theories in the history of science. However, for the false positives argument to be applicable to the success/truth case we need to make sure that such figures are genuinely available. In other words, Lewis needs to justify the two assumptions outlined in section 2, the first of which claims that 1 in 25 theories are true and the second that the false positive rate and false negative rate for success as a test for truth are both relatively low, e.g. 1 in 5. Is there is any reason to doubt the availability or accuracy of these figures for prevalence and reliability in the success/truth case?

In order to answer this question it will be useful to firstly consider how such figures are generated in the case of diagnostic testing for Alzheimer's disease. Here we have a false negative rate and false positive rate of 0.19 and 0.27 for tests based on the NINCDS-ADRDA criteria. This means that for every 100 negative test results we can expect the disease to be absent in 81 patients but present in 19 patients. Likewise for every 100 positive test results we can expect the disease to be absent in 27 patients but present in 73 patients. As Jorm (1990) notes, the only way we can arrive at these figures is to check what the test result says against information concerning the actual presence or absence of Alzheimer's disease:

> The NINCDS-ADRDA criteria state that certain diagnosis is only possible with neuropathological evidence. Thus, the clinical criteria for probable Alzheimer's

> disease can be validated if neuropathological data become available later. (Jorm 1990, p.10)

In other words, to generate figures concerning the reliability of probable diagnoses of Alzheimer's disease we need a way of detecting the presence or absence of the disease that is *independent* of what the test result says. In the Alzheimer's case this independent test is provided by neuropathological data obtained from cortical biopsies that test for the presence of clumps and tangles in the brain. Is an equivalent test available in the success/truth case?

Lewis tells us that when success is used as test for truth the false positive rate and false negative rate are both 0.2, i.e. 1 in 5. This means that for every 100 negative test results, i.e. the theory is not successful, we can expect 80 of these theories to be false but 20 of them to be true. Likewise for every 100 positive test results, i.e. the theory is successful, we can expect 80 of these theories to be true but 20 of them to be false. From the Alzheimer's case we know that these figures can only be generated if we have a way of deciding the truth or falsity of a theory that is independent of what the test results say. Unfortunately, Lewis's preferred method of determining the truth or falsity of theories does not satisfy this requirement:

> Convergent realists believe that most current theories are true for some reason, and this reason (if it is any good) distinguishes current scientific theories from past ones. The most influential realist argument is that the explanatory or predictive *success* of current scientific theories gives us good reason to think they are true; conversely, past theories are rejected because of explanatory or predictive failure. (Lewis 2001, pp.372-3)

The very possibility of providing figures for the reliability of success as a test for truth requires that we are able to distinguish between what seems to be the case (100 true theories) from what is the case (80 true theories and 20 false theories). In other words, one cannot use success as a way of identifying true theories when it is precisely this connection between truth and success that we are trying to assess. Success can either function as the test whose reliability we are concerned to assess or it can function as the independent test through which we assess another test for truth, but it cannot serve both of these functions at the same time.

Lewis does suggest another way in which we might identify true theories that does not rely on their success:

> A convergent epistemological realist believes that our scientific theories are converging on the truth, and hence that the past contains a higher proportion of false theories than the present. At a given time in the past, it may well be that false theories vastly outnumber true theories. (Lewis 2001, p.377)

If we assume, as most realists do, that this process of convergence has almost reached its limit then it follows that most current scientific theories are true (or approximately true) and that most past theories are false given that they differ from current theories in significant ways. As long as this latter judgement does not make any reference to the success or failure of past and present theories we now have a way of judging the truth or falsity of a theory in a way that is independent of the success of the theory in question. We can now generate figures for the reliability of success as a test for truth by seeing how many successful theories are true according to our present judgments of truth and falsity.

Unfortunately, there is good reason to think that this method cannot function as the kind of independent test for truth that Lewis's argument requires. We have seen that in order for the false positives argument to work we need to be able to generate figures for the reliability of success as a test for truth. In other words, we need to be able to compare positive test results (i.e. the fact that a theory is successful) with the actual presence of what we are testing for (i.e. truth). If we are to use the convergence thesis as our independent test then the claim that we have *almost* reached the limit of the convergence process is not sufficient. The fact that most current theories are true and most past theories are false does not help us to decide whether or not any particular theory is true or false. If the theory in question is a current one then the fact that most current theories are true (or approximately true) is perfectly compatible with the possibility that this particular theory is false.

To provide actual figures for the reliability of success as a test for truth Lewis needs to be able to distinguish between truth and falsity on a case-by-case basis. The only way to do this would be to assume that *all* current scientific theories are true, a claim that even the most committed scientific realist would shy away from. However, there is a far more fundamental problem with this way of providing figures for the reliability of success as a test for truth. The whole point of Lewis's argument is to defend the claim that the best explanation of the success of scientific theories is that they are true. To include the claim that science converges on the truth and the claim that all, or even some, current scientific theories are true clearly begs the question at issue.

Thus, even if Lewis could use these claims to generate reliability figures for success as a test for truth, he would still not have shown that Laudan's pessimistic induction is fallacious because this method of justifying the false positives argument assumes precisely what antirealists like Laudan want to deny.

To sum up then, neither of the two ways in which Lewis identifies true theories can function as an independent test for truth. Success cannot function as the independent test because it will produce worthless results, i.e. success will turn out to be 100% reliable as a test for truth but the circular appeal to success will make this trivially true. Similarly, the convergence thesis cannot function as the independent test because it begs the question in favour of realism. At this point, Lewis might object that these two methods were never intended to offer an independent test for the reliability of success as a test for truth. This may be true but if correct it only makes matters worse. Without a genuinely independent test for the truth of scientific theories Lewis has no way of arriving at the figures he needs to construct the false positives argument against Laudan. Lewis's failure to offer such a test seems to leave us with no option but to dismiss his probabilistic response to Laudan on the grounds that his figures are produced from nowhere.

## 6. Determining truth without success

This dismissal of the false positives argument is a little too quick. The fact that Lewis does not provide an independent test for his reliability figures should not be taken to show that no such test is available. It is still an open question as to whether or not some other method of deciding on the truth or falsity of mature scientific theories might provide the kind of independent, non-question begging test that Lewis requires. Prima facie, there are two ways in which we might attempt to do this. Firstly, we could try to determine the truth or falsity of scientific theories by identifying a property that true theories are supposed to exhibit other than success. We can call this the *indirect method* as the truth or falsity of a theory is decided by means of an intermediary property. Secondly, we could try to find a way of determining the truth or falsity of scientific theories without appealing to intermediary properties. We can call this the *direct method.* Let us examine these options in turn.

To defend success as a reliable test for truth via the indirect method requires that we identify a property of a scientific theory that is taken to show that the theory in question is true. Of course, as suggested in the previous section, if this method is to be a genuinely independent test of the reliability of success as a test for truth then the property in question had better not have anything to do with judgments concerning success. Fortunately, there are a number of other features of mature scientific theories that are generally assumed to have a strong connection to truth, e.g. explanatory power, unificatory power, simplicity, etc. Suppose then that we take one of these features as the property $\Psi$ by which we infer the truth of scientific theory T. For theory T we now have a way of driving a wedge between what our test says (i.e. true because theory T is successful or false because theory T is unsuccessful) and what is actually the case (i.e. true because theory T has property $\Psi$ or false because theory T does not have property $\Psi$). The question is now this; can we identify a plausible candidate for property $\Psi$ that will allow us to fill in the details of this argument and justify the claim that success is a reliable test for truth?

As noted above, there are plenty of possible candidates for property $\Psi$ that could serve as a way of generating figures for the reliability of success as a test for truth. However, for such figures to be acceptable we need to be able to show that a particular candidate for property $\Psi$ is indeed a good way of determining truth or falsity. The problem here is that no matter what property $\Psi$ is taken to be it will always be possible to question the inference from the presence of this property to the truth of the theory in question, e.g. there are alternative antirealist accounts of why scientific theories should have explanatory or unificatory power. So it seems that we must justify the inference from the presence of property $\Psi$ to the truth of the theory in question. In other words, we must do for our independent test exactly what it was supposed to do for the inference from success to truth. Clearly, we are now faced with an infinite regress because whatever property we appeal to for our second independent test this must be justified by a third test, which must be justified by a fourth, and so on. Therefore, the indirect method cannot be used to provide an independent test for the reliability of success as a test for truth.

There is an obvious resemblance between this argument against indirect methods of determining truth and Moore's open question argument against naturalistic definitions of 'good'. Moore (1903) argued that no matter what definition of 'good' is offered it can always be asked, 'But is it good?' because it will always make sense to ask this question of any proposed definition, Moore concludes that 'good' is a simple, non-naturalistic property that cannot be defined. To think otherwise is to commit the so-called 'naturalistic fallacy'. Like Moore's open question argument, the argument against indirect methods of determining truth suggests that no matter what property is put forward as a way of inferring the truth of a scientific theory it will always be possible to question whether or not the theory in question is true. However, the fact that this argument is similar to Moore's argument does not mean that it is open to the same sorts of objection. In fact there is an important difference between these two arguments that demonstrates precisely this point.

The argument against indirect methods of determining truth is not, in contrast to Moore's argument, about the possibility of *defining* 'true' but is rather about the possibility of justifying the claim that a particular property of a theory can be taken to show that it *is* true. Moore rules out the possibility of a defining 'good' in terms of a particular naturalistic property on the *a priori* or *intuitive* grounds that it will always make sense to ask whether or not something with that property is in fact good. But, of course, this is just to beg the question against the possibility of such definitions. In contrast, the suggestion that it is an open question whether the presence of a particular property indicates that a theory is true is an *empirical* claim based on the existence of alternative antirealist explanations of the property in question. Consequently, the stronger claim that it will *always* be an open question amounts to the well-known idea that the antirealist can account for everything the realist can (Fine 1986a; Kukla 1994). Of course, this conclusion might turn out to be false but this just shows how the argument against indirect methods differs from Moore's a priori argument against naturalistic definitions of good.

If the argument against indirect methods of determining truth is acceptable then we are forced to conclude that the only sort of independent test that will allow us to produce genuine figures for the reliability of success as a test for truth is one that can determine truth or falsity directly, i.e. without appealing to intermediate properties. Unfortunately, there is a powerful argument against the possibility of formulating any such test – this is what is sometimes called the *problem of independent access* (Giere 1988, p.109). This argument begins with the claim that our knowledge of the world must have a subjective basis, e.g. knowledge of sense impressions (Hume, Locke and Berkeley), our own existence (Descartes), paradigms (Kuhn), etc. We are then challenged to justify the claim that some theories or statements are true by virtue of their accurately representing a world that is in some sense independent of our subjective perspective. The problem for the realist is that our reliance on a particular way of looking at the world seems to rule out the possibility of gaining knowledge about the existence or nature of this representational relation. As Davidson (1984) puts it:

> Languages we will not think of as separable from souls; speaking a language is not a trait a man can lose while retaining the power of thought. So there is no chance that someone can take up a vantage point for comparing conceptual schemes by temporarily shedding his own. (Davidson 1984, p. 185)

Davidson uses this lack of a vantage point for comparing conceptual schemes to argue that the "dualism of scheme and content, of organizing system and something waiting to be organized, cannot be made intelligible and defensible" (Davidson 1984, p.189). On a similar note, philosophers like Putnam (1981) and Rorty (1991) have used the lack of a 'God's-eye point of view' or 'skyhook' to question whether the attempt to "climb outside of our own minds" (Nagel 1986, p.11) is philosophically worthwhile (see chapter 5). Although this issue is far from settled this argument would, at least at the present time, seem to rule out the *direct method* as a way of providing an independent test of the reliability of success as a test for truth.

We have grounds then for being suspicious of both the indirect and direct methods as ways of providing an independent test of the reliability of success as a test for truth. However, neither of the arguments provided above constitutes a knockdown argument against the possibility of formulating such a test. As Kukla (1994) notes, we cannot conclusively rule out the possibility of an *indirect* test for truth because it is conceivable that we may eventually be able to point to a feature of mature scientific theories whose *only* explanation is the truth of theory in question. According to this view, the reason for the current stalemate between realists and antirealists is that we have been unable to find a feature of scientific practice that one side or the other cannot explain. Likewise, although philosophy has made little headway with the problem of independent access, this in itself cannot be taken to show that there is no answer to the problem. As Nagel (1986, p.12) suggests, the fact that we have not yet provided such an answer may simply to have to with the difficulty of the task rather than its impossibility. Thus, it is also conceivable that given enough time and thought we may be able to unearth a *direct* way of determining the truth or falsity of mature scientific theories.

The fact that we cannot conclusively rule out either the indirect or direct method of determining the truth or falsity of mature scientific theories seems to suggest that we should be agnostic about the possibility of justifying Lewis's false positives argument. The validity of this argument ultimately depends on future philosophical and scientific developments that will tell us whether or not the kind of independent test it requires is actually available. However, at this point it is useful to remind ourselves of just how far we have come. Recall that the original point of the false positives argument was to provide a 'quick fix' response to Laudan's pessimistic induction. The professed advantage of this argument being that it avoided the need to tackle the details of Laudan's putative counterexamples. However, if the validity of the argument is now taken to be dependent on future philosophical developments then it seems hard to see how it provides any improvement on the usual ways of responding to Laudan. Indeed, Lewis seems to leave us worse off than the usual responses because, in order to go through, the false positives argument seems to require either that we solve (or perhaps dissolve) the problem of independent access or discover a property of scientific theories that antirealism cannot account for.

It should now be clear that Lewis's false positives argument is not a good way for the realist to respond to Laudan's pessimistic induction. In particular, it is not a good idea to make the fortunes of the explanationist defence of scientific realism parasitic on finding an independent test for truth. Let me explain why. Assume that the realist's act of faith pays off and by either the direct or indirect method we eventually get the kind of independent test the false positives argument requires. The first consequence of this development is likely to be the vindication of success as a reliable test for truth in precisely the way that Lewis outlines in the false positives argument. A second consequence of this development would surely be to eliminate the need for the explanationist defence because we now have a far more secure route to the truth that does not rely on its contingent connection to success. From the perspective of our future discovery this consequence is of no significance, it merely reflects the fact that in philosophy, just as in any form of inquiry, we find better arguments to do the same job. However, in our present predicament it is precisely this fact that shows why the false positives argument is of no use in responding to Laudan's pessimistic induction. If we cannot find an independent test for assessing the reliability of success as a test for truth then the explanationist defence is unjustified. If we can find such a test, either by the direct or indirect method, then success as a test for truth is redundant. Either way the explanationist defence does nothing to help settle the realism-antirealism issue.

## 7. Possible responses

When introducing the false positives argument I suggested that Lewis's interpretation of the explanationist defence differs from other accounts in the way it characterizes empirical success as a *reliable test* for truth. At the time, this was taken to be a harmless paraphrase of the more common characterization of truth as the *best explanation* of empirical success. However, we are now in a position to see that Lewis's characterization of the explanationist defence is no mere paraphrase. The role that success plays in Lewis's false positives argument is not the same as the role success plays in the explanationist defence. To say that success is a reliable test for truth necessarily implies that we have some way of backing up what the test claims, this follows simply from the meaning of 'reliable' and 'test'. In contrast, the explanationist defence is based on the acknowledgement that success is the *only* way we have of inferring the truth of mature scientific theories. So, by definition, the explanationist defence cannot be claiming that success is a reliable *test* for truth, at least not in the sense of 'test' that Lewis employs.

If my arguments against Lewis are correct then we seem to have reached an all too familiar conclusion, namely that the only way the realist can 'settle' the realism-antirealism issue is to beg the question in favour of his own position. Of course, it is certainly possible to envisage ways in which Lewis, or any other realist for that matter, might attempt to respond to my arguments. Most obviously, one might object to my analysis on the grounds that its central example, the Alzheimer's disease case, is unrepresentative or presented in such a way that it prejudices the case against Lewis's argument. In response to such claims all I can say is that the Alzheimer's case was chosen because accurate diagnosis of this disease can only be confirmed by autopsy. Consequently, it represents a case where there is a very clear distinction between the diagnostic test for the presence of the disease and the independent test for the reliability of this procedure. This makes it a particularly good way of illustrating the crucial flaw in the analogy that underpins Lewis's argument, i.e. the lack of an independent test for the reliability of success as a test for truth. However, I do not see how this makes the Alzheimer's example unrepresentative of the way in which diagnostic tests are evaluated for use in medical practice. My suspicion is that to be a reliable test for the presence of a disease *any* diagnostic procedure will need backing up by an independent test that confirms its accuracy. Of course, this suspicion could be wrong but it is up to the realist to prove this is so.

A more promising line of attack would be to argue for an *externalist* account of epistemological justification. In epistemology, the significance of this account lies chiefly in the way it allows one to dismiss the sceptic's demand to provide a proper *internal* foundation for our knowledge of the external world. For the epistemological externalist, such demands are unanswerable not because we have no such knowledge but rather because they rely on a faulty conception of what it is to be justified. Similarly, in the philosophy of science, externalism allows the realist to dismiss many of the sceptical arguments concerning the possibility of scientific knowledge. For example, Psillos (1999) argues that the circularity inherent in the explanationist defence will only seem vicious if one is committed to an internalist theory of justification. Psillos suggests that that once we have rejected this theory in favour of an externalist view of justification we will be able to see that the explanationist defence is acceptable and realism vindicated if abduction is, *as a matter of fact,* a reliable form of inference. This is the case even if, as the antirealist suggests, there is no non-circular way of demonstrating that abduction is reliable. From an externalist perspective, our inability to provide such a demonstration cannot be taken to show that abduction, either of the first or second-order variety, is an unjustified form of inference. In order for a rule to be justified, the externalist only requires that "the rule *is* reliable" (Psillos 1999, p.83) not that we *know* that the rule is reliable. Thus it might be possible for the realist to use this point to deflect the sort of sceptical arguments that I have presented against Lewis's false positives argument on the grounds that they misunderstand just what is required for a rule to be justified.

As Psillos notes, the move to externalism has the effect of shifting the realism-antirealism issue to more "general epistemological grounds" (Psillos 1999, p.85). It is not yet clear whether or not this shift to new ground will prove any better for realism than the old sorts of argument but it may at least enable us to escape what Wylie (1986) has called the "ascending spiral" of the realism-antirealism debate.[8] Certainly it offers the promise of new sorts of argument and may help to bridge the gap that presently exists between epistemology and the philosophy of science. In the present context, it may lead to a new form of the false positives argument that does not suffer from the sort

[8] As Kornblith (2001) notes, epistemologists are not even agreed about the precise nature of the internalism-externalism debate let alone how we might go about resolving it. For a good impression of the current state of play in the internalism-externalism debate see the collection of papers in Kornblith (2001).

of justificatory problems discussed in this paper. Whatever one thinks of this suggestion it must surely be granted that these problems are unlikely to go away so long as one continues to think that the reliability of success as a test for truth is a measurable or quantifiable property.

## 8. Conclusion

Wittgenstein (1953) once remarked that the philosopher's treatment of a question is like the treatment of an illness. However, it seems that the diagnosis of truth turns out to be an entirely different matter. There is *no* independent test for truth that will allow one to produce figures for the reliability of success as a test for truth. In the evolution of science, theories may die in our stead as Popper famously remarked but there is no possibility of performing autopsies on 'dead' theories to determine their truth status. For this reason, Lewis has not shown that Laudan's pessimistic induction is a fallacy.

# Chapter 5

# Naturalism, Realism and Scepticism

## 1. Introduction

Prior to any difficulties concerning the possibility of resolving the realism-antirealism debate, there are two far more fundamental problems facing naturalistic realism. The first problem is that naturalistic approaches to the theory of knowledge, whether in epistemology proper or the philosophy of science, attempt to explain how we know by making use of at least some of the knowledge whose epistemic status is supposed to be in question:

> The intellectual coherence of naturalism faces a graver threat...this is the charge that naturalism must ultimately be a question-begging or viciously circular philosophy of science. (Rosenberg 1996, p.23)

In the philosophy of science, this problem of *epistemic* circularity is applicable to any project that uses empirical methods or information to settle philosophical questions concerning the status of science. There is a further sort of circularity that is a result of the attempt to argue for realism from a naturalistic perspective. As both realists (Boyd 1989, p.359; Psillos 1999, p.79) and antirealists (Laudan 1996a; Fine 1986a; van Fraassen 1980) have pointed out, the explanationist defence assumes the reliability of the very form of inference whose reliability it is trying to prove, i.e. inference to the best explanation. Although this is a species of the sort of epistemic circularity that characterizes naturalism in general, it is useful to treat this problem separately so for convenience I shall refer to it as the problem of *methodological* circularity.

In the first half of this chapter I show how naturalistic realists have responded to the problems of epistemic and methodological circularity by arguing that the model of epistemology that motivates and gives both of these objections their force should no longer be taken seriously. Thus, in section 2, I show how naturalists have used a variety of developments in twentieth century epistemology, philosophy of science, and science itself, to argue that a circular approach to epistemological questions is warranted purely on the grounds that a non-circular alternative - an a priori epistemology founded on first principles - is not available. In this sense, I argue that contemporary epistemic naturalism has important links to certain anti-sceptical movements in the history of philosophy. In section 3, I argue that the same sort of anti-sceptical arguments are used to respond to the problem of methodological circularity. Here I suggest that even recent externalist treatments of this problem must be seen in terms of the rejection of traditional sceptical approaches to epistemology.

In chapters 3 and 4 we have considered some of the problems that face the realist attempt to refute Laudan's pessimistic induction. However, if the analysis presented in chapter 2 is correct, then the pessimistic induction is not the only challenge that faces naturalistic realism. In particular, we have not yet considered what naturalistic realists have to say about van Fraassen's constructive empiricism and the epistemic/metaphysical constructivism of contemporary sociology of science. In section 4 I argue that there are three types of argument used against these two forms of antirealism: philosophical, empirical and scientific. Here I suggest that the best way of arguing against these two forms of antirealism is to use scientific arguments based on evolutionary theory. However, I argue that even these arguments are unlikely to settle the issue in favour of realism. Finally, having shown some of the arguments used against their opponents, in section 5 I suggest that there is a crucial problem with the attempt to argue for realism from a naturalistic perspective. I conclude that naturalised accounts may well be better off if they did not engage in the realism-antirealism debate.

## 2. The problem of epistemic circularity

In his second discourse on method, Descartes suggests that a proper reconstruction of knowledge is possible only if we resolve:

> Never to accept anything as true that I did not know to be evidently so: that is to say, carefully to avoid precipitancy and prejudice, and to include in my judgments nothing more that what presented itself so clearly and so distinctly to my mind that I might have no occasion to place it in doubt. (Descartes 1637/1968, p.41)

Although this rule is by no means the whole of Descartes's philosophical method, it "gives the distinctive character to Descartes's investigation of knowledge, and the method which, following this rule to its limit, he uses in that investigation is famously known as the *Method of Doubt*" (Williams 1987, p.33). In fact, we might go further than this and say that the Cartesian method of doubt is what gives 'distinctive character' to most of what has passed as epistemological investigation in English-speaking philosophy for the past three centuries or so. For although post-Cartesian philosophers may have been quick to dismiss Descartes's reconstruction of knowledge based on proofs of a beneficent deity, "many have felt obliged to offer their own answers to the sceptical doubts" (Kenny 1973, pp.203-204).

As we have seen in chapter 1, the whole point of epistemological naturalism is that we explicitly rely on theories and methods that are fallible, i.e. knowledge that we may have occasion to doubt. So, naturalists are clearly guilty of the sort of 'precipitancy and prejudice' that Descartes is out to warn us against. Put simply, epistemological naturalism breaks the first and most important rule of Cartesian epistemology, the universal method of doubt. If one takes this rule seriously then we have a very powerful argument against naturalizing the philosophy of science. In using the results (strong scientistic naturalism) or methods (weak scientistic naturalism) of science to construct an epistemology of science itself, the naturalist begs the question against traditional sceptical worries about the epistemic status of science. In other words, naturalized philosophy of science is viciously circular because it makes use of the very knowledge the sceptic calls into question. For many traditional epistemologists and philosophers of science this provides sufficient ground for rejecting any attempt to naturalize the philosophy of science.

### 2.1 Quine and the problem of epistemic circularity

The *locus classicus* of the naturalistic response to the problem of epistemic circularity is Quine's (1985) 'Epistemology Naturalized'. In this paper Quine begins by suggesting that studies in the foundation of mathematics are either *conceptual* or *doctrinal*:

> The conceptual studies are concerned with meaning, the doctrinal with truth. The conceptual studies are concerned with clarifying concepts by defining them, some in terms of others. The doctrinal studies are concerned with establishing laws by proving them, some on the basis of others. (Quine 1985, p.15)

For Quine, the significance of these two types of study is the link that exists between them. Ideally one would use a conceptual study to define concepts in terms of a favoured subset of them in order that a doctrinal study might show how theorems couched in these terms will be obviously true or derivable from obvious truths.

As Quine suggests, the significance of the bifurcation between conceptual and doctrinal studies in the foundation of mathematics is that a similar distinction exists in epistemology:

> Just as mathematics is to be reduced to logic, or logic and set theory, so natural knowledge is to be based somehow on sense experience. This means justifying the notion of body in sensory terms; here is the conceptual side. And it means justifying our knowledge of truths of nature in sensory terms; here is the doctrinal side. (Quine 1985, p.16)

So, empiricist attempts to define the notion of body in sensory terms are just the epistemological equivalent of defining mathematics in terms of logic, this is the conceptual side of empiricism. Similarly, as in the reduction of mathematics to logic, the definition of the notion of body in sensory terms attempts to provide us with a proper foundation for knowledge, this is the doctrinal side.

Although empiricists like Hume had at least some success with the conceptual side of their project, Quine argues that the prospects for the doctrinal side are no better than where Hume left us. The general point here is that although empiricists have learnt the value of contextual definition, "that to explain a term we do not need to specify an object for it to refer to...we need only show, by whatever means, how to translate all the whole sentences in which the term is to be used" (Quine 1985, p.17), they have come no closer to telling us how to settle the doctrinal issue of how such definitions ground our knowledge of the world. The failure of this doctrinal project is evidenced by the failures of Russell and Carnap to reconstruct the world from sense data. For Quine, this has far reaching consequences for epistemology:

> The hopelessness of grounding natural science upon immediate experience in a firmly logical way was acknowledged. The Cartesian quest for certainty had been the remote motivation of epistemology, both on its conceptual and doctrinal side; but that quest was seen as a lost cause. To endow truths of nature with the full authority of immediate experience was as forlorn a hope as hoping to endow the truths of mathematics with the potential obviousness of elementary logic. (Quine 1985, p.18)

Just as we should give up the attempt to ground mathematics in logic so too should we give up the Cartesian attempt to ground knowledge in experience. On this view, the failure of Carnap's *Aufbau* to carry out the latter reduction is just as significant as Gödel's proof that mathematics cannot be reduced to logic.

If we must give up on the Cartesian attempt to ground knowledge in the certainty of subjective experience what is left for epistemology to do? In Quine's view we must look to the sciences, psychology in particular:

> The stimulation of his sensory receptors is all the evidence anybody has to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds? Why not settle for psychology? (Quine 1985, p.19)

Of course, the reason why we could never have settled for psychology is precisely because of the Cartesian injunction against epistemic circularity. However, as Quine says:

> Such scruples against circularity have little point once we have stopped dreaming of deducing science from observations. If we are out simply to understand the link between observation and science, we are well advised to use any available information, including that provided by the very science whose link with observation we are seeking to understand. (Quine 1985, p.19)

Here then we have Quine's answer to the problem of epistemic circularity. We need no longer worry about the circularity inherent in naturalistic approaches to epistemology simply because we have given up on the idea that a non-circular reduction of knowledge to experience is possible. In other words, giving up on the Cartesian search for certainty requires that we also put behind us the injunctions that rule out the use of science in constructing accounts of knowledge. Once we have "stopped dreaming of deducing science from sense data" (Quine 1985, p.24) we can simply ignore the problem of epistemic circularity as a relic of a bygone age, epistemology falls into place as a "chapter of psychology" (Quine 1985, p.24).

## 2.2 What is the aim of naturalized epistemology?

There is a crucial ambiguity in Quine's argument for the naturalizing of epistemology. For although Quine (1985) is clear that epistemology must become a chapter of psychology, he is less than clear about what sort of questions a suitably naturalized epistemology should answer:

> The "old" epistemology asked us how any of us knows anything at all about the world around us, and it recognized that most of what we know is based somehow on the senses. The problem was given its special philosophical character by certain facts about sense-perception, familiar from antiquity and employed to dramatic effect in Descartes's *First Meditation* and elsewhere, which seem to imply at least the possibility of the world's being quite different in general from the way it is perceived. The philosophical problem was then to explain how anyone can know that such a possibility does not obtain, and thereby know what the world is really like, not just the way it is perceived to be. Only then would the possibility of human knowledge have been explained. (Stroud 1985, p.71)

Given Quine's rejection of the Cartesian injunction against circular approaches to epistemology the most obvious question to ask is this: is naturalized epistemology an attempt to answer the traditional Cartesian problem of knowledge or do we treat this problem in the same way as the problem of epistemic circularity? To put it another way, is Quine's rejection of the "old" epistemology *wholesale* in that it rejects Cartesian questions as well as Cartesian answers or *piecemeal* in that we reject the latter but not the former?

Quine is not always clear about how we should answer these questions. As Stroud (1985) points out, there is plenty of evidence to suggest that Quine accepts at least part of the Cartesian approach to epistemology, e.g. Quine wonders how "given only the evidence of the senses, how do we arrive at our theory of the world?" (Quine 1974, p.1 quoted in Stroud 1985, p.72), similarly Quine says, "we know external things only mediately through our senses" (Quine 1960, p.1 quoted in Stroud 1985, p.72). For Stroud, such comments show that:

> Quine's conception of human knowledge and therefore of his epistemological project shares with earlier philosophers the idea of human knowledge as a combination of two quite general but distinguishable, factors – the contribution of the world and the contribution of the knowing or perceiving subject. (Stroud 1985, p.72)

In other words, although Quine rejects the injunction against circular epistemologies he fully accepts the Cartesian subject-object model of the relationship between knower and known. However, it remains to be shown that Quine also accepts the traditional problem of knowledge that is associated with this Cartesian subject-object model.

Stroud (1985) argues that when Quine (1960) talks about physical objects as "posits" or "hypotheses" he cannot be attempting to answer the traditional Cartesian problem of how we can have knowledge of the external world:

> Whether and how the physical object "hypothesis" *is* better confirmed or known is precisely what is in question when the traditional philosophical problem is raised, so the alleged superiority of the physical object "theory" cannot be taken for granted in demonstrating its superiority. Therefore, there seem to be good reasons for concluding that Quine's naturalistic epistemology does not amount to an answer to the traditional problem of our knowledge of the external world. (Stroud 1985, p.74)

In support of this reading Stroud points to Quine's (1969) claim that in terms of the doctrinal side of epistemology (the side concerning truth rather than meaning) we are no better off now than where Hume left us. So, when Quine says that the "Humean predicament is the human predicament" (Quine 1985, p.17) he means that we are no nearer a solution to the traditional problem of knowledge, whether this solution is attempted from a naturalistic or traditional perspective. However,

> The illegitimate circularity of relying on one's knowledge of nature in an attempt to "validate" that very knowledge on sensory grounds is obviously no objection to naturalistic investigations in which no such project of "validation" is in question. (Stroud 1985, p.74)

In other words, the problem of epistemic circularity can be ignored simply because it is not the purpose of naturalized epistemology to answer the question of "validation" so the question of circularity is irrelevant. So, on this reading at least, Quine's rejection of Cartesian epistemology is wholesale. Science can be used in naturalized epistemology simply because we are no longer interested in solving the traditional problem of how we can have knowledge of the external world.

There are other times when Quine seems to suggest that naturalized epistemology *is* supposed to answer the traditional problem of knowledge. Take for example his claim that because sceptical doubts about knowledge are a product of science, it is perfectly legitimate to use science itself to respond to them (see Quine 1974). On this view, Cartesian epistemology asks the right questions but mistakenly rules out the use of science in the attempt to validate our knowledge of the external world. Sceptical doubts are internal to the scientific enterprise and as such demand scientific responses:

> Science tells us that our only source of information about the external world is through the impact of light rays and molecules upon our sensory surfaces. Stimulated in these ways, we somehow evolve an elaborate and useful science. How do we do this, and why does the resulting science work so well? These are genuine questions, and no feigning of doubt is needed to appreciate them. They are scientific questions about a species of primates, and they are open to investigation in natural science, the very science whose acquisition is being investigated. (Quine 1975, p. 3 quoted in Stroud 1985, p.75)

As Stroud points out, here we seem to have Quine taking a quite different attitude to "validation" and the traditional problem of knowledge. The "Humean predicament" of not being able to "justify our knowledge of the physical world on sensory grounds" (Stroud 1985, p.74) is no longer the human predicament; it is merely the predicament of traditional epistemologists who refuse to use science in responding to sceptical doubts. On this *piecemeal* view, naturalized epistemology is the combination of Cartesian sceptical doubts and scientifically informed attempts to validate our knowledge of the external world.

**2.3 Naturalistic realism and the problem of epistemic circularity**

Contemporary naturalists have whole-heartedly endorsed Quine's response to the problem of epistemic circularity. Like Quine, prominent naturalists such as Kitcher, Giere and Boyd reject the Cartesian injunction against the inherent circularity of naturalistic accounts of science on the grounds that alternative foundationalist accounts have singularly failed to deliver the goods. For example, Giere (1985) defends his naturalistic approach to the philosophy of science on the pragmatic grounds that the foundationalist approaches of methodological foundationism (Carnap, Reichenbach and Popper) and metamethodology (Lakatos and Laudan) have both been shown to be wanting in certain crucial respects. In Giere's view, the failure of foundationalism in the philosophy of science warrants an alternative approach to these problems:

> Neither methodological foundationism nor metamethodology can break the circle or provide the norms needed to defeat relativism. This hardly proves that there is no way to achieve these ends. It does, however, provide some motivation for seeking to *understand* how a naturalized philosophy of science might fruitfully be pursued. (Giere 1985, p.339)

Here we can see that Giere justifies a naturalistic approach to the philosophy of science on the pragmatic grounds that what we have tried so far simply hasn't worked. Significantly, this includes the failure of Carnap's attempt to reconstruct the external world in sensory terms. Following Quine, Giere's tentative suggestion is that we relax the traditional injunction against circularity in the hope that we may get a better account of science.

Kitcher (1993) puts forward a similarly Quinean response to the problem of epistemic circularity:

> Skeptics who insist that we begin from *no* assumptions are inviting us to play a mug's game. Descartes's lack of success in generating an account of nature that

> would survive all possible doubt was in no way the result of deficiencies of intellect or imagination. (Kitcher 1993, p.135)

Again, the general claim is that we should reject the Cartesian demand to construct a non-circular account of knowledge on the grounds that this simply hasn't worked. However, Kitcher seems to go further than Giere in claiming that Cartesian epistemology is a 'mug's game.' This claim carries with it the suggestion that we may have been on the wrong track all along, something that is not directly implied by Giere's claim that the failure of foundationalism "hardly proves that there is no way to achieve these ends" (Giere 1988, p.339). In other words, Kitcher's argument implies that there is *no* alternative to naturalism whereas Giere's suggests that foundationalist alternatives cannot yet be ruled out.

### 2.4 Naturalistic realism and the traditional problem of knowledge

If the possibility of constructing a non-circular epistemology from no assumptions is ruled out then the way is clear to developing a 'science of science.' However, as we have already seen in the context of Quine's naturalistic epistemology, there is the crucial question of what kinds of questions a suitably naturalized philosophy of science should try to answer. On a *wholesale* view, a 'science of science' should not only reject the problem of circularity, it should also reject the traditional problem of how to justify our (scientific) knowledge of the external world. On a *piecemeal* view, a 'science of science' should, as Quine (1974) suggests, use science to answer traditional epistemological questions about how science generates knowledge about the world.

To the extent that they see naturalism as providing support for a realist account of science, naturalistic realists endorse the piecemeal view of naturalized epistemology. In other words, whilst they reject the problem of epistemic circularity on the grounds that there is no alternative to appealing to science, they accept that traditional sceptical doubts about the status of scientific theories remain. For example, Giere (1985) says:

> The general problem faced by a naturalistic philosophy of science, then, is to explain how creatures with our natural endowments manage to learn so much about the detailed structure of the world – about atoms, stars and nebulae, entropy, and genes. This problem calls for a scientific explanation. (Giere 1985, p.340)

On a similar note Boyd (1983) suggests that:

> Like the causal theorist of perception or other "naturalistic" epistemologists, the scientific realist must deny that the most basic principles of inductive inference or justification are defensible *a priori.* In a word, the scientific realist must see epistemology as an *empirical* science (Boyd 1983, p.211)

So, although they reject the Cartesian injunction against the use of science in epistemology naturalistic realists like Giere and Boyd take the main problem of epistemology to be very much the same as Descartes, namely the problem of how we can have knowledge (scientific or otherwise) of the world around us. In other words, like Quine (1974), naturalistic realists disagree with Descartes's conception of what is allowable as far as a theory of knowledge is concerned but they retain his basic problem of how we justify or arrive at knowledge of the external world. This is precisely why they are naturalistic *realists.*

The viability of a piecemeal approach to naturalized philosophy of science will be assessed shortly, but first I want to discuss the second charge of circularity facing naturalistic realism, namely the problem of methodological circularity.

### 3. The problem of methodological circularity

In order to show that the best explanation of the success of science is the approximate truth or genuine referential capacity of scientific theories the realist must assume that the success of a theory is good reason for believing it to be true. However, as a number of critics have pointed out, this last assumption is exactly what antirealists want to deny. As van Fraassen (1980) says:

> The realist asks us to choose between different hypotheses that explain the regularities in certain ways; but his opponent always wishes to choose among hypotheses of the form 'theory $T_i$ is empirically adequate'. So the realist will need his special extra premiss that every universal regularity in nature needs an explanation, before the rule will make realists of us all. And that is just the premiss that distinguishes the realist from his opponents. (van Fraassen 1980, p.21)

In a similar vein, Fine (1986b) argues that metatheoretic arguments in mathematics and philosophy must be more stringent and secure than the arguments to which they apply.[1] His problem with the explanationist defence is that it quite clearly fails to satisfy this condition:

[1] Fine (1986a) claims that arguments for realism should follow Hilbert's maxim that demands that metatheoretic consistency proofs in set theory must employ methods that are *more stringent* than those applied at the object level.

> To argue for realism, one must employ methods more stringent than those in ordinary scientific practice. In particular, one must not beg the question as to the significance of explanatory hypotheses by assuming that they carry truth as well as explanatory efficacy. (Fine 1986a, p.24)

In these two complaints we can see exactly why the explanationist defence is so controversial because it uses exactly the same form of inference at the meta-level that its opponents dispute at the scientific level. As both our critics point out, this simply begs the question in favour of the realist's belief in "the significance of explanatory hypotheses" (Fine) or a "special extra premiss that every universal regularity in nature needs an explanation" (van Fraassen). As Laudan (1996a) suggests, given the circularity of this argumentative strategy it is hard to see how the explanationist defence could convince anyone other than a committed realist. This is the problem of methodological circularity.[2]

### 3.1 Hume and the problem of induction

As Giere (1988) notes, the charge that the explanationist defence is circular "strongly resembles Hume's famous argument against the justifiability of induction" (Giere 1988, p.170) where the fact that our inductive inferences can only be justified by a further inductive inference is taken to show that no independent justification of induction is possible:

> We have said that all arguments concerning existence are founded on the relation of cause and effect; that our knowledge of that relation is derived entirely from experience; and that all our experimental conclusions proceed upon the supposition that the future will be conformable to the past. To endeavour, therefore, the proof of this last supposition by probable arguments, or arguments regarding existence,

---

[2] Laudan (1996a) has suggested that the explanationist defence is not in fact intended to convince the committed antirealist but is rather an attempt to show that realism is at least as well confirmed as any other putative explanation of science. The motivation behind this suggestion being that it licences the use of inferential principles that only realists would accept. However, as Laudan sees it, the problem with this weakening of the explanationist defence is that it although it may avoid the charge of begging the question it relies on a rather implausible account of confirmation:

> Since realism was devised to explain the success of science, it remains purely *ad hoc* with respect to that success. If realism made some novel predictions or has been subjected to carefully controlled tests, one does not learn about it from the literature of contemporary realism. (Laudan 1996a, p.135)

Laudan's point here is that in advocating the weak version of the explanationist defence the realist seems to be prepared to accept a form of *ad hoc* explanation at the philosophical level that they are quite opposed to at the scientific level.

> must be evidently going in a circle, and taking that for granted, which is the very point in question. (Hume 1748/1975, pp.35-36)

In Giere's view, it is this brand of Humean scepticism that results in the charge that the explanationist defence is circular:

> The justification of realistically interpreted hypotheses requires a principle of inference sanctioning a move from "success" to "truth." The hypothesis successfully predicts an experimental result, and one concludes (with appropriate qualifications) that it is true. Sometimes this is called "inference to the best explanation." But (said in a Humean tone of voice) what justifies this principle of inference? The almost automatic realist response is that the principle is the principle is justified by the success of science in using the principle. But this response, the antirealist replies, employs the very principle of inference at issue. The realist response, therefore, begs the question. The challenge to the realist, then, is to produce some other response. (Giere 1988, p.170)

If Giere is right about the connection between the Humean critique of induction and the charge of circularity levelled at the explanationist defence then the realist is indeed in trouble.[3] The history of philosophy is littered with many failed attempts to justify induction in the face of Humean sceptical worries. What reason do we have for thinking that attempts to justify the explanationist defence will fair any better?

It is at this point that naturalism comes to the rescue of the scientific realist. As discussed in the previous section, an important weapon in the naturalist's arsenal is the idea that we can dismiss sceptical arguments on the grounds that they tempt us into playing an unwinnable and rather pointless game. Thus, in the face of objections concerning the inherent circularity of their project, many naturalists choose to question the value of the idea that epistemology must be founded on incorrigible foundations in accordance with the Cartesian method of doubt. The naturalist's historical point is that hundreds of years of doing epistemology according to this Cartesian model have shown it to be unworkable. They argue that there is simply no alternative to appealing to empirical information in constructing accounts of our epistemic practices. It is hardly surprising that in the face of Humean calls for a non-circular justification of the explanationist defence, a naturalist like Giere is prepared to invoke a similar strategy in defence of realism:

---

[3] A similar interpretation can be found in Boyd 1983, p.213.

> The proper realist response is to reject the challenge on the grounds that it is impossible to fulfil. A realist interpretation of science requires no such ultimate justification. (Giere 1988, p.170)

Here we can see Giere defending the explanationist defence on the grounds that the Humean demands for independent justification are "impossible to fulfil" (Giere 1988, p.170). In doing so he follows the same general strategy that sees the naturalist reject the Cartesian injunction against using corrigible scientific knowledge as a basis for epistemological theorizing on the grounds that there is simply 'no alternative' available.

So, in the form of the explanationist defence, the combination of naturalism and realism requires nothing less than the rejection of the scepticism common to Descartes and Hume. However, if naturalistic realism is to be acceptable to anyone but the most committed naturalist we surely want some kind of argument to back up the idea that we should no longer take these worries seriously. After all, it does not seem unreasonable to ask why we should reject questions simply on the grounds that they seem to be impossible to answer. Where would be now if this strategy had been widely adopted in the history of science? Fortunately, there is an argument that attempts to go beyond the seemingly dogmatic claim that we should reject sceptical questions on the grounds that they cannot be answered. This argument claims that the problem of methodological circularity can be 'solved' by adopting an externalist theory of justification.

### 3.2 The externalist theory of justification

Externalist epistemologists suggest that many sceptical problems concerning the possibility of knowledge are simply a result of the particular theory of justification that lies unquestioned at the heart of Western epistemology. According to this suggestion, both sides in traditional epistemic debates rely on an *internalist* theory of justification according to which:

> Epistemology's job is to construct doxastic principle or procedure *from the inside*, from our own individual vantage point. To adopt a Kantian idiom, a [doxastic principle] must not be "heteronomous," or dictated "from without." It must be "autonomous," a law we can give to ourselves and which we have grounds for giving to ourselves. (Goldman 2001, pp.42-43)

Epistemologists like Armstrong (1973) and Goldman (2001) have dared to question the hegemony of the internalist conception by suggesting that an agent may be justified in believing p without being able to show why or how the belief in question is justified.

According to *externalist* theories of justification, an agent can be justified in believing p even when the justification for this belief is "external" to his epistemic perspective.

It is hard to overestimate just how radical the externalist theory of justification is. Indeed, for many epistemologists, to adopt externalism is simply to give up on the traditional questions of epistemology rather than provide different answers to them. This point has even been put forward as a reason in favour of externalism.[4] However, for the moment let us concentrate instead on what these answers are rather the question concerning whether or not they constitute appropriate answers to traditional epistemological questions (whatever 'appropriate' may mean here). On an externalist view, many of the trickiest problems of traditional epistemology turn out to have quite simple solutions. For example, the problem of how we can have knowledge of the external world is, for an externalist at least, simply a matter of the right sort of the right sort of relation holding between the epistemic subject and the world. Because the externalist rejects the idea that a belief or rule can only be justified if the necessary justificatory factors are cognitively accessible to the epistemic subject he does not think it important that this sort of internal justification is not available in the case of our knowledge of the external world. As a consequence of this view, the externalist can look down on traditional Cartesian discussions of this 'problem' with a mixture of amusement and confusion about what all the fuss is about.

The externalist theory of justification also makes short work of the problem of induction. For an externalist, inductive inferences are justified if, as a matter of fact, they are reliable. From this perspective, the problem with Hume's discussion of induction is that it presents us with a false dilemma. Hume argues that the only way to justify induction is either by an a priori argument concerning "relations of ideas" or an empirical argument concerning "matters of fact," this dilemma being commonly referred to as "Hume's fork." The externalist responds to this dilemma by denying Hume's assumption that induction is in need of such justification. In this sense at least, the externalist shares Rorty's (1991) pragmatist view of this problem:

> One will suffer from Hume's itch only if one has been scratching oneself with what has sometimes been called "Hume's fork" – the distinction between "relations of ideas" and "matters of fact." (Rorty 1991, p.40)

---

[4] For example, Armstrong (1973) suggests that externalist accounts of knowledge do not require the concept of justification.

Importantly, the same externalist response is also available in response to Goodman's so-call "new riddle of induction." In both cases, the externalist will simply point out that one will only take the problem of induction seriously if, like Rorty says, one has been scratching oneself with Hume's fork. Once we stop doing this we will no longer suffer from Hume's itch, the felt need to provide an internal justification of our inductive practices.

### 3.3 The externalist defence of scientific realism

As we have seen, Giere shows that there is a striking similarity between Hume's critique of induction and the sort of worries raised by Laudan (1996a), van Fraassen (1980) and Fine (1986a). It is hardly surprising then that externalism promises a similarly radical solution to the problem of methodological circularity as it does to traditional sceptical problems in epistemology. In the philosophy of science, the most recent and arguably most convincing formulation of this sort of externalist argument is due to Psillos (1999). Psillos argues that the circularity inherent in the explanationist defence will only seem vicious if one is committed to an internalist theory of justification. He suggests that that once we have rejected this theory in favour of an externalist view of justification we will be able to see that the explanationist defence is acceptable and realism vindicated if abduction is, *as a matter of fact,* a reliable form of inference. This is the case even if, as the antirealist suggests, there is no non-circular way of demonstrating that abduction is reliable. As suggested in the previous section, from an externalist perspective, our inability to provide such a demonstration cannot be taken to show that abduction, either of the first or second-order variety, is an unjustified form of inference. In order for a rule to be justified, the externalist only requires that "the rule *is* reliable" not that we *know* that the rule is reliable. Let us examine this argument in more detail.

Psillos begins by pointing that viciously circular arguments are 'premiss-circular' because "one claims to offer an argument for the truth of $\alpha$, but explicitly *presupposes* $\alpha$ in one's premises" (Psillos 1999, p.82). Following Braithwaite (1953), Psillos suggests that there is another sort of circular argument that is not 'premiss-circular':

> On the surface level, the argument is as non-circular as anything can be. It begins with the premisses [sic] $P_1,\ldots P_n$, and then, by employing an inference rule $R$, it draws a certain conclusion $Q$. However, $Q$ has a certain logical property: it asserts or implies something about the rule of inference $R$ *used* in the argument, in

> particular that *R* is reliable. Braithwaite called this argument-type 'rule circular'. (Psillos 1999, p.82)

For Psillos, the key feature of 'rule-circular' arguments is that, in contrast to 'premiss-circular' arguments, they are not obviously viciously circular because the conclusions of such arguments do not feature among the premises.

Having established the distinction between rule and premiss-circular arguments, the next step in Psillos's argument is the claim that the explanationist defence (here referred to as NMA) is, if anything, an example of a 'rule-circular' argument:

> NMA is clearly not premiss-circular. The conclusion of the meta-IBE (that theories are approximately true) is not among the premises of the argument. In fact, no assumption about the approximate truth of theories is made within the premisses, either explicitly or implicitly. (Psillos 1999, p.83)

However, even if we accept this characterization of the explanationist defence there still remains the problem of circularity, albeit in a slightly altered form:

> In a rule-circular argument one has to assume the reliability of the rule invoked in the argument. But if this assumption is based on the prior acceptance of the conclusion of the rule-circular argument, then the proponents of a rule-circular argument apparently traffic in a vicious circle. For they would have to prove the conclusion *before* they accepted the rule used to derive it. But they could not prove the conclusion unless they *first* accepted the reliability of the rule. (Psillos 1999, p.83)

Thus, the fact that the explanationist defence is rule-circular rather than premise-circular does not yet seem to have resolved the problem of methodological circularity.

It is at this point that Psillos brings externalism into the picture. Psillos claims that the rule-circularity inherent in the explanationist defence is only vicious if one accepts an *internalist* account of justification. On an internalist account, rule-circular arguments are viciously circular because in order to use an instance of a particular rule we must first show that the rule is justified, something that cannot be done without already assuming the reliability of the rule in question. Psillos suggests that we can break out of this vicious circle by rejecting the internalist demand that we must justify a rule before we can use it:

> I want to reply to this objection by denying that any assumptions about the reliability of a rule are present, either explicitly or implicitly, when an instance of

> this rule is used. Nor should the reliability of the rule be established *before* one is able to use it in a justifiable way. (Psillos 1999, p.83)

What Psillos is suggesting here is an *externalist* account of justification according to which the factors relevant to the justification of a belief or rule may be external or cognitively inaccessible to the epistemic subject. Thus, according to the externalist, one can be justified in using a rule (or holding a belief) without knowing or having reasons to believe it is reliable.

As we have seen, the significance of adopting an externalist account of justification is that it allows us to reject Humean calls for an independent justification of our inferential practices. As Psillos notes, this has important consequences for the alleged vicious circularity of the explanationist defence:

> Given externalism, all we should require of a rule-circular argument is that the rule of inference employed *be* reliable; no more and no less than in any ordinary (first-order) argument. A rule-circular argument would be no more vicious than any other first-order application of the rule involved in it. Since first-order applications are not vicious, nor is the second-order application involved in the rule-circular argument. (Psillos 1999, p.84)

So, if externalism is assumed, the critiques of Laudan (1996a), van Fraassen (1980) and Fine (1986a) fail to show that the explanationist defence is viciously circular. This is simply because the externalist refuses to accept that rules of inference must be independently justified before one can use them. All that matters to the externalist is whether or not the rule in question is reliable. This has nothing to do with any reasons we can offer for our belief but is rather an empirical fact about the reliability of the inference in question. If these *external* conditions are in place then we are quite justified in using that rule of inference to arrive at conclusions about the rule of inference itself. Hence, on an externalist view, there is nothing wrong with the explanationist defence of scientific realism because it is simply a second-order application of a reliable inferential practice.

Psillos's externalist defence of scientific realism is nothing less than a complete overhaul in the way we approach the realism-antirealism issue. Traditional discussions of this issue rest on the assumption that the realist must provide us with reasons for thinking that successful theories are true (or approximately true) rather than merely useful or empirically adequate. The explanationist defence attempts to provide such reasons by showing how the success of first-order abductive inferences is best explained by the truth of science itself. Psillos has shown that where one goes from here largely depends on the theory of justification one chooses to adopt. If one is an internalist then the appeal to a meta-abductive inference to justify a set of first-order abductive inferences is bound to seem viciously circular. However, if one is an externalist like Psillos then the use of a meta-abductive inference to justify abduction itself is only suspect if, as a matter of fact, abduction is not a reliable form of inference. In effect, this shifts the burden of proof from the realist to the antirealist for it is now the latter that must provide us with convincing reasons for thinking that abduction is unreliable.

### 3.4 Externalism and naturalism

Psillos's externalist solution to the realism-antirealism issue seems to provide a promising way for the naturalistic realist to respond to the problem of methodological circularity. Indeed, as Psillos says, the move to externalism would seem to have the advantage over internalist alternatives in that it shifts the debate away from the sort of justificatory problems that plague traditional arguments for realism:

> Proponents of [realism] have to assume an externalist theory of justification that some critics of realism might deny. But that is a different matter. That battle can be fought on general epistemological grounds which have nothing to do with the issue of circularity. (Psillos 1999, p.85)

So, if externalism is assumed, the realist can ignore the problem of methodological circularity and instead argue for their position on "general epistemological grounds." In other words, all that remains for the realist to do is show that externalist theories of justification are preferable to internalist ones. Is there any reason to think that this is any more possible than showing that realism is preferable to antirealism?

The first thing to say is that the internalism-externalism debate in epistemology is far from being settled. Indeed, as Kornblith (2001) notes, epistemologists are not even agreed about the precise nature of this debate let alone how we might go about resolving it. Part of the problem here is that externalism and internalism rely on very different conceptions of what epistemology is and what it is supposed to achieve. Thus, whether or not one finds externalism convincing is likely to depend on how one feels about traditional issues in epistemology. As one prominent internalist says:

> When viewed from the general standpoint of the western epistemological tradition, externalism represents a very radical departure. It seems safe to say that until very recent times, no serious philosopher of knowledge would have dreamed of

> suggesting that that a person's beliefs might be epistemically justified simply in virtue of facts or relations that were external to his subjective conception. Descartes, for example, would surely have been unimpressed by the suggestion that his problematic beliefs about the external world were justified if only they were in fact reliably related to the world – whether or not he had any reason for thinking this to be so. (Bonjour 2001, p.56)

As this passage suggests, internalists like Bonjour are liable to see externalism as a rather unsatisfactory way of changing the subject, an unfortunate attempt to bypass the really difficult questions concerning the justification of knowledge. On this view externalism is not simply a different way of doing epistemology, it is a repudiation of the very questions that make epistemology possible.

The fact that Bonjour (2001) mentions Descartes here is no accident for it is their different attitudes towards Cartesian epistemology that primarily separates the internalist and externalist. As Goldman points out, "internalism takes its inspiration from a perspective that has dominated epistemology since the time of Descartes" (Goldman 2001, p.36). Unlike their opponents, externalists see the Cartesian project in epistemology as a fundamentally misguided attempt to do what cannot be done. In fact, this just shows that at its most fundamental level the externalist defence of scientific realism is nothing less than the rejection of the Cartesian attempt to ground knowledge in subjective experience. In this sense, Psillos's externalist argument against the problem of methodological circularity is no different to Quine's naturalistic argument against the problem of epistemic circularity.

## 4. The evolutionary argument for scientific realism

So, the naturalistic realist defends his account from the charges of epistemic and methodological circularity by questioning the epistemological tradition on which they are based. However, even if it is legitimate to defend naturalistic realism in this way it remains to be shown that, as a matter of fact, the 'hypothesis' of scientific realism is the best explanation of the success of science. In chapter 2, I followed Rosenberg (1996) in suggesting that this turns out to be largely a matter of defeating the three main challenges to naturalistic realism: Laudan's pessimistic induction, van Fraassen's constructive empiricism, and epistemic/metaphysical constructivism. In chapters 3 and 4 we discussed the attempt to defeat the pessimistic induction. We concluded that neither the empirical approach of naturalists like Kitcher (1993) and Psillos (1999) nor the probabilistic approach of Lewis (2001) provide sufficient grounds for saying that the issues raised by Laudan can be settled in favour of (naturalistic) realism. Are naturalistic realists any more successful in their attempt to show that realism constitutes a better account of science than either van Fraassen's constructive empiricism or the epistemic/metaphysical constructivism of contemporary sociology of science?

### 4.1 Philosophical and empirical arguments against antirealism

Even if we restrict our discussion to the accounts of Kitcher, Giere and Psillos the number of arguments and criticisms aimed at constructive empiricism and epistemic/metaphysical constructivism is still quite large. In order to get a better idea of what is going on here I want to suggest that we can view these arguments in terms of three broad categories. *Philosophical* arguments against constructive empiricism and epistemic/metaphysical constructivism attempt to show that these accounts are in some way incoherent or are based on untenable philosophical assumptions. *Empirical* arguments suggest that we can reject these accounts on the grounds that they fail to adequately explain salient features of science, e.g. empirical or predictive success, the behaviour of scientists, scientific progress, etc. *Scientific* arguments are used to suggest that the results of some particular scientific theory should lead us to prefer realism to its constructivist rivals. In this section I shall concentrate on philosophical and empirical arguments

Naturalistic realists have presented a number of philosophical and empirical arguments against constructive empiricism. The most common philosophical argument involves showing that van Fraassen's distinction between observable and unobservable phenomena cannot be made in a non-arbitrary way (Kitcher 1993, pp.152-155; Giere 1988, p.128). In this sense at least naturalistic realists cover much of the same ground as analytical critiques of constructive empiricism (see Churchland and Hooker 1985). Naturalistic realists have also attempted to show that even if van Fraassen's distinction could be defended from this charge it is still possible to reject constructive empiricism on empirical or explanatory grounds. For example, Giere (1988) argues that the language and behaviour of scientists can only be explained by the existence of unobservable entities and concludes that "regarded as an empirical theory of science, van Fraassen's account is itself not empirically adequate. It fails as an empirical theory of science even when judged by its own standards" (Giere 1988, p.128). Similar strategies are used against epistemic/social constructivist accounts of science. Thus, in addition to arguing against the coherence of constructivist accounts, Kitcher and Giere follow Boyd (1983) in arguing that such accounts cannot provide adequate explanations of the success of science (Giere 1988, p.131; Kitcher 1993, pp.160-169).

It would take a proper investigation of these arguments and the evidence offered in support of them in order to decide whether or not they succeed in demonstrating the incoherence and/or empirical inadequacy of constructive empiricism and epistemic/social constructivism. However, my suspicion is that if we were to carry out such an investigation the results would prove inconclusive. Here we can make three points. Firstly, as Kukla (2000) has shown, many of the standard objections to the coherence social constructivist accounts of science have tended to underestimate the explanatory resources at their disposal. Secondly, it is not clear that van Fraassen would be unable to respond to the claim that the observable-unobservable distinction cannot be drawn in a non-arbitrary way. As Matheson (1996) says, Kitcher's attack on this distinction is "at most it is a probe, testing the boundaries of the concept of observability. It does not constitute a breakthrough for the armoured columns of the realist army" (Matheson 1996, p.466).[5] Thirdly, it is certainly not clear that one can use historical (Kitcher, Psillos) or laboratory studies (Giere) in order to settle the issues that separate the realist from their opponents. The problem is that each competing account has its own interpretation of the salient features of a particular case. In many cases this makes it difficult to see how we could even begin to compare, for example, a realist and a constructivist theory of what is going on. Is it not strange that case studies *always* end up supporting the philosophical theories they are allegedly supposed to test?

So, for the sake of argument, let us assume that philosophical and empirical arguments against constructive empiricism and social constructivism fail to establish naturalistic realism as the best explanation of science. Is that the end of the matter? Not quite. As I suggested above, some naturalistic realists have suggested that by appealing to the results of science rather than just its methods we can generate a much stronger case for scientific realism. In particular, Kitcher and Giere have argued that by appealing to the results of evolutionary theory and cognitive psychology a strong *scientific* case can be made against two of the arguments that are generally used to support antirealist accounts of science, namely the problem of independent access and

[5] Matheson (1996) provides an excellent summary of Kitcher's attempt to dispense with the three antirealist accounts discussed in chapter 3.

the argument from underdetermination. Although we have encountered them before in previous chapters let us remind ourselves of these arguments.

## 4.2 Global and local antirealism: independent access and underdetermination

The idea that there is a world independent of the way we perceive it seems to be firmly anchored in our commonsense view of the world. Of course, this cannot be taken as proof of an independent world although some realists have come very close to saying this, e.g. G.E. Moore's anti-sceptical rejoinder that he could know with absolute certainty the statement "Here is one hand, and here is another" (Moore 1939/1995). On the other hand we should be wary of any argument that dismisses the deliverances of commonsense too hastily. In particular, to reject a widely held notion like the independence of the world we should require a very strong argument to the contrary. Many philosophers think that such an argument is readily available.

In the postscript of *The Structure of Scientific Revolutions*, Kuhn (1970) contrasts his own 'puzzle solving' model of scientific progress with that "most prevalent among both philosophers of science and laymen" (Kuhn 1970, p.206):

> A scientific theory is usually felt to be better than its predecessors not only in the sense that it is a better instrument for discovering and solving problems but also because it is somehow a better representation of what nature is really like. One often hears that successive theories grow ever closer to, or approximate more and more closely to, the truth. Apparently generalizations like that refer not to the puzzle-solutions and the concrete predictions derived from a theory but rather to its ontology, to the match, that is, between the entities with which the theory populates nature and what is "really there." (Kuhn 1970, p.206)

So, from a realist perspective, Kuhn's problem-solving account lacks the "essential element" (Kuhn 1970, p.206) of a theory of scientific progress, namely the idea that successive theories are converging on the truth. However, if Kuhn is right, this is just as well:

> There is, I think, no theory-independent way to reconstruct phrases like 'really there'; the notion of a match between the ontology of a theory and its real counterpart in nature now seems to me illusive in principle" (Kuhn 1970, p.206).

Kuhn thinks this match between theories and world is "illusive in principle" because we can only ever conceptualise the latter in terms of a particular paradigm. Consequently, there is no possibility of comparing paradigms from a neutral point of view to see which of them better captures what is 'really there.'

A similar sort of argument against realism is put forward by Fine (1986a). As we have already seen, Fine argues against the explanationist defence of scientific realism on the grounds that it begs the question against the antirealist. Having done this he goes on to survey the alternative ways in which the realist might make his case:

> Hilbert's maxim applies, and we must employ patterns of argument more stringent than the usual abductive ones. What might they be? Well, the obvious candidates are patterns of induction leading to empirical generalizations. But, to frame empirical generalizations, we must first have some observable connection between observables. For realism, this must connect theories with the world by way of approximate truth. But no such connections are observable and, hence, suitable as the basis for an inductive inference. (Fine 1986a, p.24)

Although he talks about the lack of an "observable connection between observables" rather than the "match between the ontology of a theory and its real counterpart in nature" being "illusive in principle," the point Fine wants to make here is exactly the same as that of Kuhn. Both arguments "amount to the well-known idea that realism commits one to an unverifiable correspondence with the world" (Fine 1986a, p.24).

What we call this argument is of little consequence but for convenience let us follow Giere (1988, p.109) and refer to it as the 'problem of independent access' on the grounds that it questions our ability to *access* what is 'really there' in the world *independently* of what our theories say. Indeed, what matters more than nomenclature here is to appreciate just how widespread this argument is. I have chosen passages from Kuhn (1970) and Fine (1986a) to illustrate the argument but there are plenty of other formulations available. For example, the problem of independent access is an important part of 'neo-pragmatist' critiques of realism, e.g. Putnam's (1981) claim that there is no "God's-eye point of view" and Rorty's (1979, 1991) 'antirepresentationalist' critique of realism. It can also be found in the work of more analytical critiques of realism, e.g. Davidson's (1974) rejection of the scheme-content distinction as the 'third dogma of empiricism' and Dummett's verificationist arguments against realism. It is also an essential ingredient of various 'constructivist' positions, e.g. Goodman's irrealism, Latour and Woolgar (1979), etc. Many realists refer to these as global forms of antirealism.

In contrast to the global antirealist, the local antirealist accepts that there is an independent world but denies that we can have knowledge of it. For example, van Fraassen (1980) accepts that claims concerning observables may be true or false he just doesn't think we could ever have sufficient epistemic warrant for believing in the existence of such entities. As Boyd (1983) shows, this empiricist position is the result of a very simple and powerful argument that concerns the existence of empirically equivalent theories that make different claims about the unobservable parts of the world. The argument begins with the claim that, given a particular theory T, it is always possible to construct a number of alternative theories that make the same empirical predictions as T (i.e. they are empirically equivalent) but that make contradictory claims about the nature of unobservable phenomena. The problem for the realist is that:

> Since scientific evidence for or against a theory consists in the confirmation or disconfirmation of one of its observational predictions, T and each of the theories empirically equivalent to it will be equally well confirmed or disconfirmed by any possible observational evidence. Therefore no scientific evidence can bear on the question of which of these theories provides the correct account of unobservable phenomena. (Boyd 1983, pp.196-197)

Given that no scientific evidence can decide which of our empirically equivalent theories is 'right' about the unobservable phenomena, the empiricist invites us to conclude that we must make a decision to accept one of these theories for 'pragmatic' or 'instrumental' reasons. However, according to the empiricist, we are never entitled to make this decision on the grounds that the theory in question is true simply because we have no way of deciding when this might be the case.

### 4.3 From rats to realism

Recall from chapter 2 that Giere's constructive realism is based on the idea that the best way to conceive the relationship between theories and the world is in terms of the notion of similarity rather than truth. This claim is important because it allows Giere to sidestep certain problems concerning the notion of truth:

> Most theories of science, whether old or new, assume that any representational relationship between theory and reality would have to be understood as a "correspondence" between scientific statements and the world. The fate of any understanding of theories as somehow representing reality has thus been linked to the fortunes of a correspondence theory of truth. It is here that the battle is usually joined. The interpretation I have offered above undercuts these arguments by

> denying the common assumption. There is, on this account, no direct relationship between sets of statements and the real world. (Giere 1988, p.82)

And again:

> In recent years many philosophical differences between realist and anti-realist interpretations of science have been formulated in semantical terms like 'truth' and 'reference'. In what sense, if any, can a scientific hypothesis be said to be "true"? Do theoretical terms genuinely "refer"? My interpretation bypasses these semantical questions and focuses directly on the respects and degrees of claimed similarity between models and real systems. (Giere 1988, p.93)

So, according to Giere, we can solve or perhaps even ignore many of the traditional problems to do with the relationship between theory and world simply by thinking of it in terms of *similarity* between model and real system rather than a *correspondence* between statement and fact. For example, Giere claims that his non-linguistic account can deal with many of the objections to the notion of approximate truth because it reveals what approximate truth talk is really about, namely the respects and degrees of similarity between model and world. On this view, worries concerning how theories can be 'approximately true' are simply a result of thinking in the wrong way about the relationship between theory and world, i.e. as a relationship between language and world, rather than non-linguistic model and world.

Giere's move from truth to similarity is a clever way of avoiding certain problems that are connected with the correspondence theory of truth. However, one might wonder whether Giere has not just replaced the problematic relation of correspondence with the equally problematic relation of similarity. Of course, by suggesting that a model is always similar to a real system in certain respects and degrees Giere can avoid the most obvious objection to the similarity relation, namely Wittgenstein's point that everything is similar to everything else. However, this is not the only way in which we might object to the similarity relation.

Firstly, as Giere points out, the move to similarity carries with it the threat of relativism because the truth or falsity of theoretical hypotheses crucially depends on how we choose to specify of appropriate respects and degrees of similarity. The suggestion is that this specification must be the result of a process of social sanction and agreement "which are totally independent of our models and of how the world really is" (Giere 1988, p.108). Secondly, although the introduction of the similarity relation sidesteps issues concerning how we can establish correspondence between statements and facts, it does not address the more general issue of how anything can represent anything else. In other words, because the similarity relation relies on drawing an intelligible division between models and real systems, it is still open to Kuhn's claim that there is no theory-independent way of distinguishing between what our theories say and what is 'really there.' In other words, the move from correspondence to similarity does not get rid of the problem of independent access.

It should come as no surprise that Giere responds to both of these arguments by appealing to his naturalism. Having argued that both are rooted in traditional empiricist epistemology, Giere outlines his naturalistic responses to these two arguments. Here is how he responds to the argument that judgments of similarity are purely social constructs:

> Traditional empiricism is particularly vulnerable to an attack based on post-Darwinian biology. The effect of evolution on our sensory apparatus is known to have been particularly strong. Animals are capable of incredibly fine discriminations among objects in their environment without benefit of social conventions. And so – being fairly intelligent, talking primates – are we. (Giere 1988, p.109)

The suggestion here is that, like our animal cousins, we arrive at certain discriminations between objects *naturally* as a result of our evolutionary history. Giere wants to say that this is also very likely to be the case with our judgments of respects and degrees of similarity. On this evolutionary view, our shared cognitive and sensory apparatus are at least as responsible as social forces for widespread agreement on such issues. Therefore, the conclusion of complete social determination does not follow.

Giere offers a similarly naturalistic response to the problem of independent access:

> Approaching the problem from the perspective of the cognitive sciences and evolutionary naturalism allows one to bypass several centuries of fruitless philosophical debate. Rats (Tolman 1948; O'Keefe and Nadel 1978), and even wasps (Gallistel 1980, 345-49), have the capacity to construct internal "maps" of their environment. They produce those maps through causal interaction with the world in a way that yields useful similarities with that world. Of course, evolution produced the neural capacity for generating such maps – again as a result of long-term causal interactions with the world. And versions of those same mechanisms exist as well in the human brain (P.S. Churchland). (Giere 1988, p.110)

Again, Giere's suggestion is that an appeal to cognitive science and evolutionary biology will free us from the mistakes of the past. In this case we need only recognize that, like rats and wasps, our cognitive apparatus is simply the result of a "long-term causal interaction with the world" that works "in a way that yields useful similarities with that world" (ibid. p.110). Here we have nothing less than an evolutionary defence of scientific realism whereby the evolutionary process guarantees that our judgments hook up with the world in a representationally significant way. In other words, from an evolutionary perspective we don't need the 'view from nowhere' to make sure that our judgments and theories answer to a world beyond our ken, our status as products of the evolutionary process guarantees that this is the case.

### 4.4 Rehabilitating truth

As a fellow naturalistic realist, Kitcher (1993) is just as aware as Giere of the problems that face the 'commonsense' idea of correspondence between language and world. However, unlike Giere, Kitcher does not see the need for a surrogate relationship (i.e. similarity) to replace the problematic notion of correspondence. Indeed, Kitcher is quite prepared to defend the commonsensical notion of truth as correspondence:

> Semantic facts concern the relation between language users and nature. In virtue of the state of the language user and the state of the rest of the world, there is sometimes a relation – the relation of reference – between the words spoken or written and items in the world. In consequence, the statement represents the world as being in some particular way. The statement is true just in case the way in which the world is represented is the way it really is (Kitcher 1993, p.128)

A more straightforward definition of the correspondence theory of truth one could not wish to find, but how does Kitcher propose to defend this account of truth in the face of the problem of independent access?

Kitcher's understanding of the problem of independent access is largely based on Kuhn's claim, discussed earlier, that the match between the ontology of a theory and its 'real' counterpart in nature is "illusive in principle." In response to Kuhn's worries about this correspondence or match between theory and world, Kitcher argues that Kuhn's account of progress itself seems to require the notion of truth:

> Why does he contend that later theories in mechanics (both Newton's and Einstein's) improve on their predecessors as "instruments for puzzle-solving"? Could it be that, at the level of individual statements, Kuhn is prepared to permit talk of truth and of approach to the truth? If so, how does such discourse avoid

> presupposing a match that is "illusive in principle"? If not, what exactly *is* a puzzle solution (since it cannot be a *true* answer to a question) and what does it mean for puzzling solving to improve? (Kitcher 1993, p.130)

Even if Kitcher is right to criticize Kuhn's argument against any kind of match between theories and world it is hard to see how these sorts of consideration help Kitcher defend the notion of truth as correspondence. If the account of progress as puzzle solving does let truth in through the back door this only shows that Kuhn, like Kitcher, relies on the notion of correspondence between statements and the world. This doesn't dispense with the problem of independent access; it merely suggests that Kuhn's account is inconsistent with his pronouncements on truth.

So, Kitcher's critique of the Kuhnian account of puzzle solving will not do as a response to the problem of independent access, he still needs to tell us why the notion of correspondence is not "illusive in principle." Another way of defending this relation is suggested in the following passage:

> The correspondence theory of truth is often held to involve extravagant metaphysics, but, I claim, its roots lie in our everyday practices. We explain and predict the behavior of our fellows by attributing to them states with propositional content. We explain and predict the differential successes of our fellows in coping with the world by supposing that there are relations between the elements of their representations and independent objects. (Kitcher 1993, p.131)

Although it is not made quite as explicit, Kitcher follows Giere here in defending the correspondence relation on the naturalistic grounds that everyday and scientific accounts of human behaviour presuppose this kind of relation to exist. Of course, this presupposition could be mistaken but Kitcher thinks there is a very good argument that rules out this possibility. Thus, the passage quoted above continues with the following argument:

> Those with correct beliefs about spatial relations can navigate their way more successfully than those who have faulty beliefs, and they can do so because their beliefs correspond to the ways in which the constituents of the local environment are arranged. (Kitcher 1993, p.131)

What we have here is nothing less than an evolutionary argument for the correspondence theory of truth. So, like Giere, Kitcher thinks that scientific realism can be justified on evolutionary grounds.

## 5. The incompatibility of naturalism and realism

So, if Kitcher and Giere are right, the results of evolutionary theory and cognitive psychology give us strong *scientific* reasons to reject two of the main arguments used in support of antirealism. Of course, this scientific defence of realism requires us to accept the use of science in dealing with sceptical objections but as we have already seen this is simply a consequence of rejecting the Cartesian injunction against epistemic circularity. Is this right? If we accept the results of science will we have to admit that realism is right after all?

It seems clear that neither van Fraassen nor his constructivist cousins in the sociology of science would accept that a move to naturalism necessarily lends support to realism. Indeed, as we saw in chapter 2, van Fraassen thinks that evolutionary theory can be used to provide support for constructive empiricism, not realism. Similarly, many sociologists argue that although we should be realists about the social world, we should be antirealists about the natural world (Fuller 1988, Collins 1985). As a consequence, they are likely to suggest that we should put more weight on their own sociological explanations rather than the alleged 'results' of evolutionary theory. So, as with the philosophical and empirical arguments used against van Fraassen and epistemic/metaphysical constructivism, I think we should say that the scientific arguments produced in defence of realism are likely to prove inconclusive. However, as I will now argue, there is a far more worrying conclusion in store for the naturalistic realist. If I am right then naturalism may not even be compatible with realism.

### 5.1 The received view of antirealism

Papineau attempts to capture various antirealist positions in terms of a distinction between 'US-style' antirealism and 'Dummett-style' antirealism. According to Papineau (1996b), to be realist about any body of knowledge involves the conjunction of two theses:

1. The independence thesis – "our judgments answer for their truth to a world which exists independently of our awareness of it."

2. The epistemic thesis – "by and large, we can know which of these judgments is true."

(Papineau 1996b, p.2).

So, one cannot be a realist if one denies either the *metaphysical* claim that the world is independent of our judgments or the *epistemic* claim that we can have knowledge concerning which of these judgments is true. For example, Devitt (1991) has referred to a view he calls 'fig-leaf realism', a position that accepts the claim that the world "exists independently of our awareness of it" but denies that we can "know which of these judgments is true." For Papineau, there is no point in calling such a view realist, even in a minimal 'fig-leaf' sense, simply because the knowledge thesis is a necessary condition of being realist about any particular body of knowledge.

On this view, any position that denies either the independence thesis or the knowledge thesis is a form of antirealism. This allows Papineau to make an important distinction between two quite different forms of antirealism:

> There are two traditional alternatives to realism, defined by their rejection of one of these theses. The *idealist or verificationist* tradition abandons the independence thesis, arguing that the very notion of some further world, beyond the world as we perceive it, is incoherent. *Sceptics,* by contrast, abandon the knowledge thesis, and accept that we cannot know the truth about the world. (Papineau 1996b, p.3)

According to this analysis, *all* forms of antirealism fall into one of two camps. Idealists deny the claim that there is an independent world but accept the claim that we can know which of our judgments are true. In this camp Papineau places traditional idealism, verificationism, and phenomenalism because they attempt "to *uphold* our claims to knowledge, by arguing that such claims should not be read as answering to a world beyond our ken" (Papineau 1996b, p.5). In contrast, sceptics accept the claim there is an independent world but deny that we can know the truth about that world. In this camp Papineau places van Fraassen's constructive empiricism, fictionalism and instrumentalism because they all want to "*reject* any scientific claims to knowledge about the unobservable world, precisely on the grounds that such claims *do* answer to a world beyond our ken" (Papineau 1996b, p.5). To distinguish these two forms of antirealism Papineau refers to the first of these positions as 'Dummett-style antirealism' and the second as 'US-style antirealism.'

In a moment we shall have cause to question this bipartite classification of antirealists as either US-style sceptics or Dummett-style idealists. Having said this we must grant that Papineau's account is not without its merits. For example, one cannot deny the claim that only confusion can come from a failure to distinguish between the restricted doubts of the former and the more global 'idealist' doubts of the latter. As we have already seen, an empiricist like van Fraassen wants to say that only claims about observable entities can be known to be true or false, claims about unobservable entities being merely 'empirically adequate.' As Papineau suggests, this is an epistemic dispute concerning what sorts of judgments we can know to be true, not a dispute concerning how those judgments relate to the world. For example, van Fraassen does not want to say that we can conclusively rule out the existence of the unobservable entities postulated in successful scientific theories but rather that we can never have warrant to believe in their existence. In contrast, the idealist takes issue with realism and US-style antirealism on the grounds that both rely on the idea that it makes sense to think about the world independently of the way we describe it. In other words, the Dummett-style antirealist denies that this sort of independence is possible and, in many cases, may not even be desirable. On this view, US-style antirealism is "just a quaint form of late Platonism" (Rorty 1991, p.52), an unfortunate consequence of taking the observable-unobservable distinction too seriously.

### 5.2 The therapeutic critique of realism

This idea that one can characterise antirealism in terms of a broad distinction between empiricist accounts that dispute our knowledge of unobservables and idealist accounts that deny the independence of the world is a popular one. Indeed, something like this distinction can be found in a number of recent attempts to classify and explain the varieties of antirealism currently available in the philosophy of science (Boyd 1983; Fine 1986a; Kitcher 1993; Rosenberg 1996). However, as it stands, this classification will not do because it fails to capture an important difference between two different 'antirealist' positions. These two positions correspond to two quite different senses in which we can deny the realist claim that reality or the world is independent of our awareness of it.

Recall from chapter 2 the claim made by many sociologists of science that the world (or perhaps our idea of the world) is 'constructed' in the sense that facts about reality are the consequence rather than the cause of scientific inquiry (Latour and Woolgar 1979). According to such accounts our judgments literally construct or decide what sorts of entities populate the world. Compare this constructivist account with the argument put forward by philosophers like Davidson (1984), Rorty (1991, 1998), Fine (1986a, 1986b, 1991) and Putnam (1981) that just as we should deny the claim that the world is independent of our judgments of it, so too should we reject the opposite conclusion that what is real or exists is somehow caused by or a result of the judgments or decisions we make. According to these 'antirealist' accounts, realism and epistemic/social constructivism are equally unfortunate attempts to do something that cannot be done. The question concerning the independence/dependence of the world is a bad one that we should refuse to answer. For the purposes of this chapter I shall refer to this argument as the *therapeutic critique of realism.*

So, although they both deny that the world is independent of our awareness of it there is clearly an important difference between constructivist accounts of science and therapeutic critiques of realism. The constructivist objects to the independence thesis on the grounds that he has an alternative account of how our judgments relate to the world. In contrast, a proponent of the therapeutic critique has no such positive account to offer. The problem with Papineau's classification and others like it is that it singularly fails to capture this difference between constructivist and therapeutic accounts by classifying both as examples of 'Dummett-style antirealism.' In my view, much damage comes from such simplistic classifications because the therapeutic critique presents a quite different challenge to realism than that posed by the idealist-sounding accounts of contemporary social constructivism. Even if we do not think it is possible or desirable to reject the sorts of questions that divide realists and antirealists we should at least acknowledge the fact that attempts to do this are not the same as accounts that have something positive to say about this issue.

### 5.3 Why should naturalists be realists?

At the heart of the therapeutic critique of realism is the idea that we should reject the kinds of questions and distinctions that support and give meaning to the dispute between realists and antirealists. For example, Fine (1986a, 1986b) suggests that we should reject the idea common to realism and antirealism that science is in need of an interpretation (see chapter 6); Rorty (1991, 1998) argues that we should reject the notion that language or ideas *represent* a world that is independent of our judgments (see chapter 7); and Putnam (1981) urges to reject the distinction between "the dichotomy between objective and subjective views of truth and reason" (Putnam 1981, p.ix). Although they offer quite different diagnoses of the problem, all of these critiques suggest that we will only be able to free ourselves of the idea that we should be either realist or antirealist by rejecting some cherished parts of the Western philosophical tradition. Haven't we seen this sort of argument before?

In response to the problem of epistemic circularity we have seen naturalistic realists arguing that we can ignore this problem on the grounds that it is based on an outdated Cartesian injunction against the use of empirical information in answering the sceptic. Similarly, in response to the problem of methodological circularity we have seen naturalistic realists attempt to defend the circularity of the explanationist defence by arguing that we can ignore this problem on the grounds that there is no alternative (Giere) or that it relies on an outdated model of epistemic justification (Psillos). In both cases naturalistic realists follow the same strategy as proponents of the therapeutic critique realism:

1. Demonstrate that a particular philosophical problem cannot be, or has not been, solved.

2. Suggest that the reason for our failure to provide a solution to the problem is that it is founded on outdated or mistaken ideas about the topic in question.

The problem for the naturalistic realist is that although they can use these two strategies to respond to the problems of epistemic and methodological circularity the very same strategies can be used to suggest that we ought to get rid of realism as well. In other words, the naturalistic realist needs to tell us why the use of these strategies is appropriate against problems of circularity but not against realism.

What are the options here? Given that the realism-antirealism issue has not yet been solved it seems that the naturalistic realist cannot object to the first step of the therapeutic critique. So, if they are to resist the therapeutic critique then naturalistic realists must deny the second step in this argument. In other words, they must deny the claim that the realism-antirealism issue rests on outdated or mistaken ideas. However, the reason philosophers like Rorty and Putnam reject the realism-antirealism issue is because it seems to rely on the outdated conception of the lone Cartesian subject who must reconstruct his knowledge of the external world from his own internal perspective. The problem for the naturalistic realist is that this seems to be precisely the same Cartesian model that the naturalistic realist objects to when it is used to support the problems of epistemic and methodological circularity. So, in order to support their naturalism they must reject Cartesian epistemology but, if the therapeutic critique is correct, they must also rely on it in order to support their realism. If this is so how can we accept a position that both rejects and relies on a particular set of epistemological assumptions?

This is a difficult question but it does not quite show that naturalism and realism are incompatible. The naturalistic realist might be able to justify a piecemeal rejection of Cartesian epistemology according to which we can retain the Cartesian subject-object model of the epistemic subject but reject the injunction against circularity. In this way they could follow Quine (1985) in giving non-Cartesian answers (naturalism) to Cartesian questions (realism). However, the problem with this suggestion is that philosophically it is going to be very difficult to argue for due to the very close connection that exists between the method of universal doubt and the Cartesian model of the epistemic subject. Arguably the very reason that we have such a model is precisely because the universal method of doubt forces us to begin our epistemological investigations with incorrigible knowledge of our own existence. In other words, it is the demand for certainty that makes us think that there is a problem concerning our knowledge of the external world. The problem for naturalistic realism is that it is far from clear that if we drop this demand the subject-object model of inquiry (and its related sceptical questions) will remain intact. Certainly we at least need an argument to suggest that it would.

## 5.4 Can Darwin help?

If I am right about the connection between the demand for certainty and the subject-object model of investigation then the naturalistic realist needs to show that one can reject the former without compromising the latter. As far as I am aware no naturalistic realist has recognised this difficulty with their approach let alone provided an argument in defence of it. Perhaps though we can anticipate some likely responses. When they are really in trouble one can usually rely on the naturalistic realist to appeal to science in order to defend their position. Thus, it is more than likely that they would pursue a similar strategy in response to my objection. So, let us imagine that the naturalistic realist attempts to defend their differentiated attitude to Cartesian epistemology by suggesting that the subject-object model (but not the demand for certainty) is supported by the results of a particular scientific theory. Which theory might be used to show this?

Given the arguments discussed in section 4 it is most likely that the naturalistic realist would select evolutionary theory for this task. Unfortunately, there are two important reasons why such an appeal will not provide the kind of justification that is required. Firstly, as both Kitcher (1993) and Rosenberg (1996) have noted, our understanding of the effect of our evolutionary history on our relationship to the world is not advanced enough for it to settle philosophical debates about representation. Secondly, even if we do allow the appeal to evolutionary theory, it is not clear that it supports the subject-object model of epistemology. Indeed, there is good reason to think that our evolutionary history shows such models to be defunct given that there is no sense in which we could ever be out of touch with our environment. Indeed, this is precisely why Davidson (1984) and Rorty (1991) frequently invoke evolutionary stories in order to support their rejection of realism.

## 6. Conclusion

If the preceding arguments are correct then it seems that we have good reason to be suspicious of attempts to argue for realism from a naturalistic perspective, at least without further argument. My suggestion is that the unresolved tension that lies at the heart of naturalistic realism coupled with its failure to refute its main antirealist opponents gives us prima facie reason to drop the commitment to realism in naturalised philosophy of science. Instead, we should follow those proponents of the therapeutic critique and attempt to find a position that is neither realist nor antirealist. Whether or not this is possible will be the focus of the remaining two chapters.

# Chapter 6

# Fine and the Four Problems of NOA

## 1. Introduction

If we are to follow up on the idea that naturalized philosophy of science might be better off without having to take sides on the realism-antirealism issue then we need to offer an alternative account of the ontological and epistemic status of the entities and processes postulated by modern science that is in some sense neither realist nor antirealist. If the analysis of previous chapters is correct then this means that we need an account of science that avoids such claims as "science corresponds to the truth," "evolutionary theory is empirically adequate," and "electrons are just social constructions." Fortunately, there is an obvious place to begin our investigation into the possibility of developing an alternative to realist and antirealist interpretations of science. In relation to his own naturalized account of science, Kitcher (1993, p.134) suggests that it *might* be possible to get everything his account needs from Arthur Fine's (1986a) "natural ontological attitude" (henceforth NOA). Although I shall not be directly concerned with pursuing this suggestion in relation to Kitcher's account, I want to follow Kitcher in thinking that the best place to start the attempt to outline a nonrealist position in the philosophy of science is with Fine's NOA.

Indeed, there are two important reasons why it is desirable to begin our investigation with this view. Firstly, NOA is perhaps the most developed attempt to outline a philosophy of science that is neither realist nor antirealist. If there are any major obstacles to being a nonrealist in the philosophy of science then we are more likely to encounter them in a well-thought out account like Fine's than in half-baked alternatives. Secondly, NOA is also the most famous attempt to develop a nonrealist philosophy of science. As a consequence, many of the problems that affect Fine's NOA have been taken to be problems that affect the nonrealist project as a whole. Consequently, if one wants to defend such a position in the philosophy of science then one had better have something to say about how these problems can be solved or at least why they can be ignored.

In the following section I begin by showing how NOA is Fine's answer to the perceived failures of realist and antirealist explanations of science. I show how Fine views such explanations as unnecessary attachments to a commonsensical view of science and the world around us. I conclude by showing how this commonsense view of science and philosophy ultimately constitutes a rejection of the commonly held view among philosophers that science is in need of an interpretation or theory. If section 2 attempts to put NOA in a positive light, then section 3 draws attention to its flaws and inconsistencies. In this section I argue that a detailed examination of the many critiques of NOA suggest that it faces four main problems: the problems of acceptance, collapse, asymmetry and closure. My claim is that any attempt to resurrect or rehabilitate nonrealism in the philosophy of science must either resolve these problems or show how they can be avoided. Here I suggest that the best place to turn in seeking such solutions might lie in the "neo-pragmatism" of Rorty and Putnam. The problem with this suggestion is that Fine has argued at length that neither Rorty nor Putnam can provide an alternative to realism and antirealism because all they can offer is just another form of antirealism. In section 4, I outline Fine's argument for this claim and suggest that it is based on a misunderstanding of what Rorty and Putnam are trying to say. Finally, in section 5, I use this conclusion to suggest that a pragmatic version of NOA may not after all be out of the question.

## 2. What is NOA?

Fine's argument for NOA can be broken down into four parts. Firstly, Fine aims to show that realism is not a viable philosophical position. Secondly, he argues that alternative antirealist accounts of science are just as unacceptable, though for quite different reasons. Thirdly, Fine presents us with an analysis of the realism-antirealism debate in terms of the 'core position', a commitment to the results of science that both realists and antirealists can accept. Finally, on the basis of this analysis, Fine puts forward the *natural ontological attitude* as a mediating position between realism and antirealism. Let us begin with the case against realism.

### 2.1 The case against realism

Along with Laudan and van Fraassen, Fine (1986a, 1986b, 1991) has been one of the most vociferous and outspoken opponents of realism in recent times. Thus, at the beginning of his seminal discussion, Fine (1986a) confidently pronounces the death of realism (Fine 1986a, p.112) and suggests that the only remaining task is to convince the supporters of realism that their subject has passed on. Of course, as Fine realises, many realists are reluctant to accept this conclusion:

> To be sure, some recent philosophical literature has appeared to pump up the ghostly shell and to give it new life. I think these efforts will be seen and understood as the first stage in the process of mourning, the stage of denial. But I think we shall pass through this stage and into that of acceptance, for realism is well and truly dead, and we have work to get on with, in identifying a suitable successor. (Fine 1986a, p.112)

We shall see shortly what Fine thinks this "suitable successor" is. However, before we can move from the "the process of mourning" to the stage of acceptance we need to know why attempts "to pump up the ghostly shell" are unsuccessful. What justification does Fine have for saying that realism is dead?

Fine has two major arguments against the claim that realism is the best explanation of the success of science. Firstly, he argues that the explanationist defence is viciously circular because it relies on the very inference whose reliability it aims to prove. He concludes that to argue for realism "we must employ patterns of argument more stringent than the usual abductive ones" (ibid. p.115). Secondly, he argues that the only way of producing such an argument is blocked by "the well-known idea that realism commits one to an unverifiable correspondence with the world" (ibid. p.116). These two arguments are what we have been calling the problem of methodological circularity and the problem of independent access respectively. As we have discussed these two arguments in the previous chapter I do not want go into too much detail here. However, we should note that for Fine they show that "the strategy of arguments for realism as a good explanatory hypothesis...*cannot* (logically speaking) be effective for an open-minded nonbeliever" (ibid. p.116).

Although the problems of methodological circularity and independent access constitute Fine's major reasons for rejecting realism, he does present two further arguments. The first of these arguments concerns the problem of explaining the 'small handful' - the fact that only a small handful of very similar scientific theories are ever considered as possible successors in a given field at any particular time. As Fine says, this feature of scientific theorizing raises three distinct questions:

1. Why only a small handful out of the (theoretically) infinite number of possibilities?
2. Why the conservative family resemblance between members of the handful?
3. Why does the strategy of narrowing the choices in this way work so well?

Of course, the realist will answer these questions by suggesting that already existing theories are approximately true so we should only consider other theories whose ontologies and laws match them, hence the small handful and the family resemblance. Further, if these already existing theories are approximately true then so too will be the restricted set of alternatives, hence the reason why the narrowing down of possibilities works so well.

Perhaps unsurprisingly, Fine objects to each of these realist responses. Firstly, he suggests that the response to the first question is inadequate because even if we restrict successor theories to those that resemble already existing ones we still leave more than a 'small handful'. Secondly, Fine points out that in order to explain the family resemblance between members of the small handful the realist must assume that confirmation is a mark of approximate truth. The problem with this claim is that there is plenty of evidence against the inference from confirmation to approximate truth, e.g. Laudan's pessimistic induction, but none in favour of it simply because we have "no independent evidence for the relation of approximate truth itself" (ibid. p.118), i.e. we are back to the problem of independent access. Thirdly, although the realist can explain why the strategy of the small handful works in relation to the ground already covered by existing theories, the real difficulty is "rather in how to account for the successes of the later theories in new ground, or with respect to novel predictions, or in overcoming the anomalies of the earlier theories" (ibid. p.118). Here Fine says that the realist has no special resources with which to do this because "nothing in the approximate truth of the old theory can guarantee (or even make it likely) that modifying the theory in its less-confirmed parts will produce a progressive shift" (ibid. p.119).[1]

So, according to Fine, realism cannot adequately explain the problem of the 'small handful'. He reaches a similar conclusion in relation to "another methodological favourite of the realist" (ibid. p.120) - the 'problem of conjunctions.' This is the

[1] In addition, Fine argues that instrumentalism is a better theory than realism because in response to the three questions concerning the small handful "the instrumentalist scores two out of three, whereas the

problem of explaining why the conjunction of two well-confirmed theories should turn out to be a reliable predictive instrument. Again, the realist wants to say that the reason this happens is that both well-confirmed theories are approximately true and so their conjunction is also approximately true, which in turn explains its predictive success. Fine responds to this argument by again pointing out that the realist must again make the question-begging move from explanatory success to approximate truth. Further, the realist 'explanation' also mistreats the concept of approximate truth because "the tightness of an approximation dissipates as we pile on further approximations" (ibid. p.121). Accordingly, we should expect the conjunction of two approximately true theories to be less reliable than either on its own, but as Fine says, "this is neither what we expect nor what we find" (ibid. p.121). Again, Fine's conclusion is that realism fails to explain a feature of science that its proponents had assumed could *only* be explained by realism.

### 2.2 The case against antirealism

So, like van Fraassen and Laudan, Fine is quite adamant that the explanationist defence of realism cannot settle the realism-antirealism debate. However, in contrast to van Fraassen and Laudan who are both committed antirealists, Fine makes it quite clear that the "death of realism" should not lead us to endorse some form of antirealism because "just as realism will not do...neither will antirealism" (ibid. p.136). Of course, given the various ways in which "antirealism" is used and understood, we need to know precisely what sorts of position Fine has in mind here. What is "antirealism" according to Fine?

For Fine, all forms of antirealism can be classified in terms of how they react to the four basic components of scientific realism:

1. A belief in a definite world structure.
2. A belief in the possibility of substantial epistemic access to that structure.
3. The belief that science (and, to some extent, achieves) all the epistemic access to the definite world structure that realism holds to be possible.
4. A standard, model-theoretic, correspondence theory of truth; where the model is just the definite world structure posited by realism and where correspondence is understood as a relation that reaches right out to touch the world.

realist, left to his own devices, has struck out" (Fine 1986a, p.120). For more details of this claim see Fine 1986a, pp.119-120.

The first two theses here correspond quite closely to the metaphysical and epistemological theses of Papineau's definition of realism (Papineau 1996b).[2] The third thesis concerning the aims of science is, according to Fine, what makes realism "scientific". The final thesis is intended to capture the semantic aspect of realist accounts of the correspondence between theory and world.

According to Fine, all forms of antirealism reject the metaphysical, epistemological and scientific components of realism.[3] In addition, they reject the semantic component, the idea of truth as some sort of correspondence between theory and world. However, Fine argues that in rejecting this semantic component, antirealists typically "divide among themselves over the question of whether or not that realist picture of truth ought to be replaced by some other picture" (Fine 1986b, p.137). Those antirealists that do attempt such a replacement are *truthmongers*, a type of antirealist who promotes a consensus or pragmatic theory of truth according to which truth is defined in terms of the consensus of a particular community or perhaps in terms of "what it is good to believe" (James 1907/1981). For Fine, the most recent examples of truthmongering antirealism can be found in Putnam (1981), Rorty (1979) and Kuhn (1970). As we shall deal with Fine's arguments against the 'truthmongers' in a later section let us now turn to his discussion of the other main type of antirealist, the *empiricists.*

For Fine, empiricists are antirealists who deny the four theses of realism but do not attempt to replace the correspondence theory of truth with an alternative conception of what truth is. Perhaps unsurprisingly, Fine suggests that the paradigmatic example of this type of antirealist is van Fraassen (1980) who, in contrast to 'truthmongers' like Putnam and Rorty, is not after an alternative definition of truth as coherence or acceptance. Indeed, as we have already seen, van Fraassen argues for a literal construal of truth according to which theoretical claims concerning unobservables are capable of being true or false. Of course, the whole point of van Fraassen's constructive empiricism is to deny that we can ever *know* whether or not such claims are true or false. However, unlike the truthmongers, he does not think we need a different theory of truth to show that when *properly* construed such claims can be interpreted as true or

[2] As we shall see in the next chapter this similarity explains Fine's inability to outline an acceptable nonrealist alternative to realism and antirealism.

false. As Fine says, van Fraassen "seeks to impose limits on our epistemic attitudes" (Fine 1986a, p.143) rather than to alter our conception of truth. Indeed, one could argue that in proposing isomorphism between non-linguistic models and the observable world, van Fraasen does not even reject the correspondence theory of truth but is rather out to question its epistemic scope. No matter. For the sake of argument let us allow Fine's interpretation of van Fraassen's constructive empiricism. What is so wrong with this position?

Fine suggests that any attempt to restrict knowledge to the observable world must face two sorts of question:

1. Can the boundary be marked off in a way that does not involve suspicious or obnoxious assumptions?
2. What is the rationale is for putting the boundary just there, and to what extent is that placement arbitrary?

With regard to the first question Fine suggests that the empiricist is in trouble as soon as we ask whether or not the property of "being observable" is itself observable. Van Fraassen (1980) claims that what counts as observable must be a result of science. The problem with this requirement is that it is not clear that the property of "being observable" will itself count as observable according to science. As Fine says, this is "surely some thing forced on us a priori by this empiricist philosophical stance" (Fine 1986a, p.144). So, pace van Fraassen, it is philosophy rather than science that decides what is to count as observable.[4] Fine argues that the only alternative is to not ask questions concerning the observability of the property of "being observable". However, given that this manoeuvre is a completely ad hoc restriction on the questions we may legitimately ask, this is just to say that empiricism cannot avoid obnoxious assumptions concerning the limits it imposes on our epistemic attitudes.

What about the second question concerning the rationale for placing the boundary of observability in a particular place? Here Fine joins other critics of van

---

[3] It is not clear whether or not Fine thinks that there are forms of antirealism that do not reject all three of these theses. See Fine 1986a, p.137.

[4] Fine shows that the same conclusion can be reached by considering how the empiricist might try to determine that something was observable by appealing to scientific theories. The problem, as Fine sees it, is that in using such theories the empiricist will want restrict belief to the observable aspects of the theory. However, since observability was what we were trying to establish it seems that we are stuck unless we put an a priori restriction on what counts as observable.

Fraassen in arguing that there is no good reason for restricting belief to observables. However, unlike many critics who focus on the fuzziness of the observable-unobservable distinction, Fine grants that a sensible distinction can be drawn here but demands to know why we should restrict belief to observables rather than to what is unobservable but detectable. Again, Fine argues that van Fraassen (or, for that matter, anyone else who wants to restrict belief to observables) has no non-arbitrary way of showing why inquiries into unobservable yet detectable entities should not warrant belief:

> Surely the end product of such inquiries, when each one pursues a specific area of uncertainty or possible error, can only be a very compelling scientific documentation of the grounds for believing that we are, actually, detecting atoms. Faced with such substantial reasons for believing that we are detecting atoms, what, except purely a priori and arbitrary conventions, could possibly dictate the empiricist conclusion that, nevertheless, we are unwarranted actually to engage in *belief* about atoms? (Fine 1986a, p.146)

Like Hacking (1983), Fine thinks that the evidence in favour of the entities detected by modern science gives us very strong reason to believe that they exist. On this view, the *only* rationale for restricting belief to that which we can directly observe is to place arbitrary a priori constraints on what sort of things we are allowed to believe in; constraints that are no part of science itself. Fine invites us to conclude that in trying to avoid the metaphysics of realism, empiricists like van Fraassen commit the "sin of epistemology" (Fine 1986a, p.147).

### 2.3 The 'core position' of realism-antirealism

Many antirealists share Fine's criticisms of the explanationist defence of realism. Similarly, many realists share Fine's distrust of 'truthmongering' attempts to define truth in terms of coherence or the empiricist attempt to restrict our knowledge to the observable realm. As a consequence, Fine turns out to be an ally and an enemy to both realists *and* antirealists in the philosophy of science. However, this does not mean that he does not appreciate the sort of intuitions that lie behind these competing positions. Indeed, Fine claims that there is a very simple yet powerful argument that motivates the realist:

> I suggest that a more simple and homely sort of argument is what grips him. It is this, and I will put it in the first person. I certainly trust the evidence of my senses, on the whole, with regard to the existence and features of everyday objects. And I

> have similar confidence in the system of "check, double-check, check, triple-check" of scientific investigation, as well as other safeguards built into the institutions of science. So, if the scientists tell me that there really are molecules, and atoms, and $\psi$/J particles, and, who knows, maybe even quarks, then so be it. I trust them and, thus, must accept that there really are such things with their antecedent properties and relations. (Fine 1986a, p.127)

This 'homely line' is intended to capture what 'really' lies behind the idea that we should be realists about science. For Fine, it is this homely argument rather than a "sophisticated form of abductive argument" (Fine 1986a, p.126), i.e. the explanationist defence, that leads realists to argue against various antirealist attempts to devalue or downgrade the status of scientific claims.

Fine's suggestion that the homely line is the chief motivating factor behind realism is itself compelling. Indeed, many realists would probably accept Fine's homely line as a fair characterization of the 'heart' of their position with the possible caveat that it is not the *only* motivation. However, it is less clear what the antirealist would have to say about the homely line. Initially, Fine seems to suggest that the homely line is *not* an argument the antirealist would endorse. For example, Fine argues that another way of understanding the homely line is in terms of how it relates to antirealist interpretations of unobservable entities:

> Moreover, if the instrumentalist (or some other member of the species of the species "nonrealistica") comes along to say that these entities and their attendants are just fictions (or the like), then I see no more reason to believe him than to believe that *he* is a fiction, made up (somehow) to do a job on me; which I do not believe. It seems, then, that I had better be a realist. (Fine 1986a, p.127)

Here then it seems that the homely line is unavailable to the antirealist, at least if the antirealist in question belongs to the species "nonrealistica" (whatever this might mean). However, at other times, Fine insists that the homely line must be available to both realist *and* antirealist:

> Now, do you think that Bohr, the archenemy of realism, could toe the homely line? Could Bohr, fighting for the sake of science (against Einstein's realism) have felt compelled either to give up the results of science, or else to assign its "truths" to some category different from the truths of everyday life? It seems unlikely. And thus, unless we uncharitably think Bohr inconsistent on this basic issue, we might well come to question whether there is any necessary connection moving us from accepting the results of science as true to being a realist. (Fine 1986a, pp.127-128)

The first thing to say is that Fine seems to have conflated two separate issues here. There is surely a difference between giving up the results of science and assigning the "truths" of science to a category different from the truths of everyday life. Can't one do the latter without necessarily doing the former? Isn't this precisely what Bohr's nonrealist Copenhagen interpretation tries to do?

For the moment let us postpone this issue until later sections and allow that Fine's attribution of the homely line to Bohr makes sense. The really controversial aspect of Fine's analysis is the claim that the homely line is acceptable to all (or perhaps just most) antirealists:

> Let me use the term "antirealist" to refer to any of the many different specific enemies of realism: the idealist, the instrumentalist, the phenomenalist, the empiricist (constructive or not), the conventionalist, the constructivist, the pragmatist, and so forth. Then it seems to me that both the realist and the antirealist must toe what I have been calling "the homely line." That is, they must both accept the certified results of science as on a par with more homely and familiar claims. (Fine 1986a, p.128)

This passage is the key to understanding Fine's analysis of the realism-antirealism debate for he claims that at its core is a shared commitment to the homely line. On this view, both the realist and the antirealist accept the results of science in the same way as they accept more homely truths. Fine refers to this shared acceptance of scientific truths as the *core position* of the realism-antirealism debate. As we shall see there is much debate concerning the soundness of this claim.

If the homely line is the core position of their dispute then the difference between realists and antirealists must lie in what they add to this core. Indeed, this is precisely what Fine suggests:

> The antirealist may add onto the core position a particular analysis of the concept of truth, as in the pragmatic and instrumentalist and conventionalist conceptions of truth. Or the antirealist may add on a special analysis of concepts, as in idealism, constructivism, phenomenalism, and in some varieties of empiricism. These addenda will then issue in a special meaning, say, for existence statements. Or the antirealist may add on certain methodological strictures, pointing a wary finger at some particular inferential tool, or constructing his own account for some particular aspects of science (e.g., explanations or laws). (Fine 1986a, pp.128-129)

Similarly, the realist adds to the core position:

> A desk-thumping, foot-stamping shout of "Really!" So, when the realist and the antirealist agree, say, that there really are electrons and that they really carry a unit negative charge and really do have a small mass (of about $9.1 \times 10^{-28}$ grams), what the realist wants to add is the emphasis that all this is really so. "There really are electrons, really!" (Fine 1986a, p.129)

So, according to Fine, both the realist and the antirealist accept the results of science, it is just that the antirealist wants to add the qualifier "Not Really!" and the realist responds with his desk-thumping, foot-stamping shout of "Really!" On this view, the realism-antirealism debate takes on a pantomime-like quality where the protagonists are engaged in a pointless shouting match over what Nagel has called "preferred modes of speech" (Nagel 1961, p.152).

### 2.4 The natural ontological attitude

Having argued that realism and antirealism are merely 'additions' to the core position, Fine suggests that we are presented with the possibility of a third, mediating alternative:

> It is the core position itself, *and all by itself.* If I am correct in thinking that, at heart, the grip of realism only extends to the homely connection of everyday truths, and that good sense dictates our acceptance of the one on the same basis as our acceptance of the other, then the homely line makes the core position, all by itself, a compelling one, one that we ought to take to heart. (Fine 1986a, pp.129-130)

This is the *natural ontological attitude* (NOA), the decision to accept the core position on its own without any additional realist or antirealist embellishments. For Fine, this is the best way of developing a "commonsense epistemology" of science (Fine 1986a, p.130).

Fine advertises NOA as having several advantages over its realist and antirealist rivals. Firstly, and most controversially, he claims that NOA allows us to treat truth "in the usual referential way" (ibid. p.130) and licences "ordinary referential semantics" (ibid. p.130). Secondly, he claims that it does this without committing us to an overly strict 'progressivism' according to which science takes us closer and closer to the truth. According to Fine, NOA is perfectly compatible with a Kuhnian account of paradigm shifts and wholesale changes of reference, "it sanctions reference and existence claims, but it does not force the history of science into prefit molds" (ibid. p.131). Thirdly, NOA allows us to make sense of the realist's claim that electrons *really* exist without running into the problem of independent access. As Fine says:

> I think the problem that makes the realist want to stamp his feet, shouting "Really!" (and invoking the external world) has to do with the stance the realist tries to take vis-à-vis the game of science. The realist, as it were, tries to stand outside the arena watching the ongoing game and then tries to judge (from this external point of view) what the point is. It is, he says, *about* some area external to the game. The realist, I think, is fooling himself. For he cannot (really!) stand outside the arena, nor can he survey some area off the playing field and mark it out as what the game is about. (Fine 1986a, p.131)

Here we can see Fine making the same point that we examined in an earlier chapter, namely that there is no external point of view or skyhook from which you can compare how we talk or think about electrons with how electrons *really* are. However, Fine is now able to present us with a quite radical solution to this problem of independent access. He suggests that "we can stand off from the electron game and survey its claims, methods, predictive success, and so forth" (ibid. p.131) even though no external point of view is available. This is simply because for Fine "we are in the world, both physically and conceptually. That is we are among the objects of science, and the concepts and procedures that we use to make judgments of subject matter and correct application are themselves part of that same scientific world" (ibid. p.132).

Of course, from a realist perspective, this 'internal' solution to the problem of independent access is no solution at all because it doesn't address the main problem of how our judgments answer to a world beyond our ken. However, this is precisely why Fine thinks NOA is better than realism, namely because it acknowledges the impossibility of 'solving' this problem. In other words, what realists see as a flaw Fine regards as a virtue. This last point demonstrates just how far Fine is prepared to go in outlining his alternative to realism and antirealism. In effect, he is prepared to question the very purpose of philosophy of science itself by denying that science is an activity in need of an external interpretation or justification. He expresses the point of view in terms of the following metaphor:

> The realisms and antirealisms seem to treat science as a sort of grand performance, a play or opera, whose production requires interpretation and direction. They argue among themselves as to whose "reading" is best. I have been trying to suggest that if science is a performance, then it is one where the audience and crew play as well. Directions for interpretation are also part of the act. If there are questions and conjectures about the meaning of this or that, or its purpose, then there is room for those in the production too. The script, moreover, is never finished, and no past dialogue can fix future action. Such a performance is not susceptible to a reading or interpretation in any global sense, and it picks out its own interpretations, locally, as it goes along. (Fine 1986a, p.148)

We can now see why the pantomime metaphor introduced in an earlier section was quite apt. In Fine's view, realists and antirealists are as much a part of the play of science as the scientific "actors" whose actions and ideas they attempt to interpret. The irony of the situation is that realists and antirealists have failed to realize that the "audience and crew play as well"; philosophy of science is just a play within a play

If realism and antirealism are just "idle overlays" (ibid. p.149) to science, unnecessary interpretations that seek to place science in a larger context, what (if anything) is left to say about science from a philosophical perspective? Fine suggests that we should resist the temptation to say *anything* general about what science is or what it should be. Instead, Fine argues that we should "let science speak for itself" (ibid. p.150) because we "cannot actually do more, with regard to existence claims, than follow scientific practice" (ibid. p.132). Whether or not such a minimalist approach to the philosophy of science even warrants the title "philosophy of science" is highly debatable. As Leplin says, it is arguable that NOA is "not an alternative to realism and antirealism, but a pre-emption of philosophy altogether, at least at the metalevel" (Leplin 1997, p.174). Leplin certainly has a point here. If the final court of appeal for ontological and epistemological issues raised by science is composed of scientists themselves, what is left for the philosopher of science to do let alone say? If seems that if NOA were to be universally adopted Fine would not only have signed the death warrant of realism, he would also have guaranteed the death of his subject.

We will return to this issue concerning the consequences of adopting NOA in later chapters. However, I now want to focus on the various problems that face Fine's account because, if correct, they show that there is very little reason to think that anyone could or ever would embrace NOA as a genuine alternative to realism and antirealism.

## 3. What is wrong with NOA?

NOA has not had quite the revolutionary effect its author had hoped for.[5] Unlike its biblical namesake, Fine's NOA is not without its faults. Almost two decades after the publication of his seminal paper, the realism-antirealism debate in the philosophy of science goes on unabated.[6] One reason for the failure of Fine's project is that realists (Musgrave 1996, Devitt 1991, Kitcher 1993, Psillos 1999) and antirealists (van Fraassen

[5] See Fine 1986a, p.134.

[6] For an excellent recent survey of the current healthy state of realism see Psillos 2000.

1985) alike have failed to see just how NOA is supposed to differ from their own positions:

> Fine's own 'natural attitude' to science is not unproblematic. A point that has been repeatedly made is that Fine's NOA is inherently unstable: under close inspection, it collapses into realism or its rivals. (Psillos 1999, p.228)

> I must confess to finding NOA elusive: in his attacks on realism, Fine seems to become an antirealist, and in his rejection of antirealism, he appears to become a realist. (Kitcher 1993, p.134)

For this and other reasons it is fair to say that most philosophers of science no longer consider NOA to be a viable way of slipping between the horns of realism and antirealism. Rumours of the 'death of realism' may have been exaggerated but it seems that NOA is truly dead and buried.

In order to establish NOA as a genuine mediating position between realism and antirealism Fine (1986a) relies on the claim that both the realist and antirealist can accept the 'homely line.' According to Fine, this means that whatever else the realist and antirealist may disagree about they are at least united in thinking that our attitude toward the truth-status of scientific claims must not substantially differ from our attitude toward the truth-status of more mundane, everyday claims. As we have seen, this is what Fine calls the 'core position' of the realism-antirealism debate. The "Natural Ontological Attitude" is simply to accept the 'core position' without adding any additional realist or antirealist interpretations. There are a number of problems with this line of argument, two of which regularly feature in discussions of Fine's work.

### 3.1 The problem of acceptance

The first problem with Fine's account concerns the claim that both realists and antirealists can accept the so-called 'core position.' As Musgrave (1996) points out, this claim seems to be at odds with many of the positions that are traditionally referred to as antirealist:

> This is mysterious. As usually understood, the realism-anti-realism issue centres precisely on the question of truth. As usually understood, realists can accept Fine's core position, but anti-realists cannot. Positivists deny the existence of the 'theoretical entities' of science, and think that any theory which asserts the existence of such entities is *false*. Instrumentalists think that scientific theories are tools or rules which are *neither true nor false*. Epistemological anti-realists like van Fraassen or Laudan concede that theories have truth-values, even that some of them might be true, but insist that no theory should be *accepted as true*. None of

> these anti-realist positions, as usually understood, is consistent with Fine's core position. (Musgrave 1996, p.45)

Here then we seem to have at least three kinds of antirealist who it seems will not be able to accept Fine's 'core position'. If to accept the core position means that an antirealist must "believe in the existence of those entities to which his theories refer" (Fine 1986a, p.130), most (if not all) antirealists will be unable to accept it. It seems then that Fine's NOA cannot be the core position of the realism-antirealism debate.[7]

The majority of Fine's critics have generally taken this problem to show that NOA is unworkable. However, in an introductory paper to various issues in the epistemology of science, Papineau (1996) attempts to explain a way in which we might preserve Fine's claim that the homely line is the core position. Papineau suggests that the dispute between Fine and Musgrave is actually a result of their addressing different issues. He begins by pointing out that the term 'antirealism' is ambiguous because it often stands for two quite different positions:

> Dummett's 'antirealism', like more traditional idealism and verificationism, seeks to *uphold* our claims to knowledge, by arguing that such claims should not be read as answering to a world beyond our ken. By contrast, American 'antirealism' wants to *reject* any scientific claims to knowledge about the unobservable world, precisely on the grounds that such claims *do* answer to a world beyond our ken. (Papineau 1996b, p.5)

Papineau's suggestion is that Fine only intended his claim concerning the 'homely line' to apply to the debate between the realist and the Dummett-style antirealist. However, because Musgrave is more interested in the quite separate debate between the realist and the American-style antirealist, he mistakenly thinks that the 'homely line' is supposed to apply here as well.

At first glance this appears to be a very neat way of explaining the apparent flaws in Fine's account. However, on closer inspection, it becomes clear that no matter how stylish Papineau's explanation might be, it cannot be right. For, as Papineau himself seems to realise, it is simply not true to say that Fine is not interested in the debate between the realist and the American-style antirealist. Throughout Fine's NOA papers there are numerous discussions of this form of antirealism. Indeed, as we shall see in section 3, Fine devotes an entire paper to discussing the merits of instrumentalism as an alternative to realism. If instrumentalism, the attempt to draw a "sharp distinction

---

[7] Devitt (1991) and Psillos (1999) have put forward similar arguments.

between concepts applicable to observational situations and theoretical concepts" (Musgrave 1996, p.147), doesn't count as a form of American-style antirealism then nothing does. Thus I think we must conclude that Papineau's attempt to save NOA from critics like Musgrave fails. This leaves us with our original problem:

**Problem 1:** How can the 'homely line' be the 'core position' of the realism-antirealism dispute when it is clearly inconsistent with positivist forms of antirealism?

### 3.2 The problem of collapse

We have established that Fine's analysis of the realism-antirealism debate in terms of the 'core position' is suspect because it is inconsistent with positivist (or 'American') forms of antirealism. However, as Musgrave (1996) and Psillos (1999) have noted, this problem appears almost insignificant when considered alongside Fine's views on truth:

> When NOA counsels us to accept the results of science as true, I take it that we are to treat truth in the usual referential way, so that a sentence (or statement) is true just in case the entities referred to stand in the referred-to-relations. Thus, NOA sanctions ordinary referential semantics, and commits us, via truth, to the existence of the individuals, properties, relations, processes, and so forth referred to by the scientific statements that we accept as true. (Fine 1986a, p.130)

> NOA recognizes in 'truth' a concept already in use and agrees to abide by the standard rules of usage. These rules involve a Davidsonian-Tarskian referential semantics, and they support a thoroughly classical logic of inference. (Fine 1986a, p.133)

These are probably the most quoted passages in all of Fine's work and they certainly constitute his most problematic statements. For once we understand NOA to be committed to understanding truth 'in the usual referential way' or as sanctioning 'ordinary referential semantics' it is difficult to see exactly how it is supposed to differ from standard forms of realism. After all, isn't the claim that 'a sentence (or statement) is true just in case the entities referred to stand in the referred-to-relations' just another way of saying that 'truth is correspondence with reality'? It seems that, contrary to Fine's intentions, NOA collapses into just another form of correspondence-style realism.[8]

---

[8] It also shows that NOA is a position that is unacceptable not only to positivist antirealists but also to Dummett-style antirealists.

One philosopher who sees Fine's comments on reference as an unfortunate blip in his attempt to get beyond the realism-antirealism issue is Richard Rorty. Rorty (2003) suggests that we see Fine's appeal to 'ordinary referential semantics' as a consequence of his attachment to the notion of ontological commitment, which Rorty suggests he could well do without. In order to support this claim Rorty picks up on a number of comments in Fine's work that seem to go against the need for such realist notions:

> NOA, Fine says, "tries to let science speak for itself, and it trusts in our native ability to get the message without having to rely on metaphysical or epistemological hearing aids". So why, I am tempted to ask Fine, would you want to drag in a semiotic hearing aid such as "ordinary referential semantics"? (Rorty 2003, p.5)

Why indeed? For Rorty, such passages show that Fine's comments on reference, seized on so effectively by Musgrave and Psillos, betray "a confusion between existential commitment on the one hand and a profession of satisfaction with a way of speaking or a social practice on the other" (ibid. p.6). According to Rorty, Fine would have been better off if he had been satisfied with scientific ways of speaking (i.e. 'let science speak for itself') without feeling the need to back up this intuition with the kind of existential commitments provided by a robust theory of reference.

So, once again, we have a plausible account of how Fine might avoid the charge of misunderstanding the realism-antirealism issue. In this case, Rorty suggests that Fine can make NOA distinct from the realism of Musgrave and Devitt by simply dropping the notion of ontological commitment. The implicit assumption being that such notions do not form an integral part of Fine's overall project – that they are somehow peripheral. However, reflection on the details of Fine's account reveals at least two reasons why this assumption cannot be correct. Firstly, Fine's allusion to truth 'in the usual referential way' and other such notions are characteristic of his entire output, not mere isolated remarks. Secondly, the very name Fine chose for his project (the "natural *ontological* attitude") suggests that the notion of ontological commitment is not something he could easily give up. Of course, neither of these factors shows that Rorty's proposed makeover of NOA is not possible, or even undesirable, but they do suggest that any changes we might make to NOA are likely to be more than purely cosmetic. We will return to this issue in the next chapter but for the moment we are left with the following problem:

**Problem 2:** Is it possible to use terms like 'in the usual referential way' and 'ordinary referential semantics' in a way that does not automatically lead to realism?

### 3.3 The problem of asymmetry

Because of their ubiquity in the extensive critical literature on Fine and the fact that their conjunction is generally taken to show that Fine's NOA is not a genuine alternative to the realism-antirealism debate, the first two problems I have discussed constitute what I shall call the 'received view' of NOA. However, as I will now show, there are two other important problems facing Fine's account that have received considerably less critical attention. The first of these has to do with Fine's attempt to establish the following claim:

> Metatheorem 1: If the phenomena to be explained are not realist-laden, then to every good realist explanation there corresponds a better instrumentalist one. (Fine 1986b, p.154)

Fine's (1986b) argument for this 'Metatheorem' is based on the claim that all realist explanations of scientific practice must employ instrumental reliability as an intermediate connection between the truth of a theory and its success in practice. Fine's claim is that the account obtained by focusing on this connection (i.e. the instrumentalist one) is 'better' than the realist account because no further explanatory work is done by talking about truth rather than reliability. Kukla (1994) argues that there are two main problems with this argument. Firstly, it is wrong because it is simply not true to say that all realist accounts are mediated by instrumental reliability. Secondly, and more importantly for our purposes, it seems to be inconsistent with Fine's earlier claim that neither realism nor antirealism provides a better explanation of science, i.e. that their failure is symmetrical. I will concentrate on the second of these issues.

For Kukla, the failure of Fine's Metatheorem argument ultimately lies in the fact that *neither* side in the realism-antirealism dispute has identified a practice that is essential to science that their opponent cannot account for. [9] This results in what we might call Kukla's symmetry principle:

[9] Perhaps unsurprisingly, Kukla claims that the antirealist's attempts to find antirealist-laden practices have been just as unsuccessful as the realist's attempts to find realist-laden practices.

> So far as is known at present, every scientific practice is compatible with both realism and antirealism. (Kukla 1994, p.965)

The problem is that as long as this principle holds, and no new practices are introduced, any argument that attempts to privilege either realism or antirealism will inevitably beg the question. So, Fine's claim that instrumentalist explanations are better than realist ones cannot be based on the claim that the former can explain something about science that the latter cannot. Rather this claim must have something to do with the character of the instrumentalist explanations themselves. What are the options here? The only one available seems to be the claim that, even though the realist can account for everything the instrumentalist can, the latter's explanations are preferable on the grounds that they provide a logically weaker account of the phenomena. But, of course, this claim is precisely what the realist disputes, i.e. the realist does not think that logical weakness or parsimony is a feature of good explanations. Thus, the only way the instrumentalist can argue that his way is better is to beg the question against the realist.

Thus, it seems that Fine's Metatheorem argument begs the question against the realist. At this point we can recall that the charge of 'question-begging' is the foundation of Fine's arguments against both the explanationist defence of realism and van Fraassen's defence of instrumentalism.[10] It is surprising to find then that Fine actually proposes an argument that commits the same question-begging manoeuvre! Indeed, as Kukla shows, it is also strange that Fine should have an argument in favor of instrumentalism given his earlier claim that neither realism nor antirealism would do as our 'philosophy of choice':

> I agree with Fine that there's no comfort for realists in the miracle argument, or in any other arguments that simply presuppose that some non-empirical virtue like explanation has epistemic significance. In the end, Fine also maintains that there is no case for antirealism. This is another assessment with which I concur. However, it is difficult to reconcile this assessment with Metatheorem 1. How can one say that instrumentalism 'has no argument' ([1986], p.168) if it is true that 'to every good realist explanation there corresponds a better instrumentalist one' ([1986], p.154)? (Kukla 1994, p.971)

What is going on here? Should we take the Metatheorem argument to be a repudiation of Fine's earlier views or is there a way in which we can see it as being supplementary

---

[10] Indeed, we can also recall that Fine seems to save some of his more vitriolic criticisms of realism for errors of this kind.

to them? In order to clear up this confusion Kukla suggests that we view the situation in the following way:

> I suppose that the claim that instrumentalism has a better case is merely provisional, to be withdrawn when it is realized that the instrumentalist account simply begs a different question. Be that as it may, Fine must be faulted for not making the provisional nature of the conclusion, and its subsequent withdrawal, explicit. (Kukla 1994, p.971)

Kukla's suggestion is that, although the Metatheorem argument appears to support instrumentalism, its real purpose is to undermine realism after which point the rejection of instrumentalism will follow *on other grounds*.

Broadly speaking, I think that this way of interpreting Fine's Metatheorem argument is correct, i.e. it is an additional step in his attack against the realism-antirealism debate as a whole. Certainly, it clears him from the charge that his treatment of realism and antirealism is asymmetrical. However, the real puzzle about the Metatheorem argument is the role it is supposed to play in Fine's overall project. This additional feature of Fine's attack on realism is particularly difficult to understand once we realize that it relies on the very same question begging maneuvers that he finds so unconvincing elsewhere. What we need then is an explanation this seemingly inconsistent position. The problem with Kukla's suggestion is that, although it shows that the Metatheorem argument is not inconsistent with his attempt to get beyond the realism-antirealism issue, it tells us virtually nothing about how it relates to the rest of Fine's project.[11] So, again, we are forced back to our original problem:

**Problem 3:** Why do we need the Metatheorem argument, or something like it, to dispense with instrumentalism and allied forms of antirealism?

### 3.4 The problem of closure

In discussing the problem of asymmetry we saw that Kukla was able to reject Fine's Metatheorem argument on the grounds that it could not establish the required asymmetry between realist and instrumentalist accounts of science. However, for Kukla, this failure is not merely restricted to this argument but in fact applies to all known attempts to privilege antirealism over realism or vice versa. For even if the Metatheorem

[11] Of course, I am not thinking here of sociological or political motivations but rather of what it is about Fine's overall project that motivated the introduction of the Metatheorem argument.

argument is rejected and symmetry restored to Fine's account Kukla still sees it as an open question as to whether or not Fine's original arguments against the realism-antirealism issue justify its rejection:

> Fine's arguments against realism and antirealism are not so definitive as to undermine all hope of finding an advantage for one side or the other. (Kukla 1994, p.972)

Thus, unlike Fine, Kukla refuses to give up on the possibility that one side in the realism-antirealism debate may eventually come out on top.

A careful examination of Kukla's discussion reveals that his reluctance to give up on the realism-antirealism issue primarily results from the following two claims:

> 1. [Fine's] discussion does not rule out the possibility that there are essential practices which are realist-laden (or antirealist-laden for that matter).
>
> 2. Even if we are persuaded that there cannot be a telling argument from scientific practice, it is still possible that some argument unrelated to scientific practice might do the job.
>
> (Kukla 1994, p.972)

The first claim alludes to the possibility of discovering a practice whose success is crucially dependent on the scientist/s in question holding realist/antirealist views. The ability of such a case to settle the realism-antirealism dispute would then hang on establishing the truth of the counterfactual 'If scientist $x$ had not held realist/antirealist views then practice $y$ would not have been successful'. The thinking behind Kukla's second claim is less obvious for it seems to rely on the somewhat vague intuition that there might be an argumentative strategy that realists and antirealists have failed to consider. The trouble for Fine is that if either or both of these possibilities are genuine then his rejection of the realism-antirealism dispute is premature.

On the surface, Kukla's first claim seems convincing but is this really the case? Kukla seems to think that it must be because if certain epistemic practices (involving the adoption of realist or antirealist beliefs) are presupposed by a particular scientific practice then, by definition, they cannot be explained in terms of an alternative framework.[12] Now as an argument against Fine's Metatheorem argument this is all well and good because it shows why realist and instrumentalist accounts are not inter-translatable in the way Fine suggests. The problem is that Kukla wants to draw a much more ambitious conclusion, namely that we cannot rule out a resolution to the realism-antirealism dispute because there may be examples of realist/antirealist-laden practices that will do this for us. However, this is a rather strange claim precisely because *we already know of the existence of both realist-laden and antirealist-laden scientific practices but neither has lead to a resolution of the realism-antirealism dispute.* Indeed, the reason that this is the case is because of an important distinction between realism-antirealism as a *descriptive* issue that affects the day-to-day working lives of scientists (e.g. Is this result an artefact of my procedures or a genuine entity?) and realism-antirealism as a *normative* philosophical project (e.g. How do we justify realism/antirealism as our philosophy of choice?). The existence of this distinction is precisely why Fine can acknowledge the existence of both antirealist and realist-laden practices (a descriptive issue) and still think that we would be better off without philosophical discussions of realism-antirealism (a normative issue). The problem with Kukla's first objection is that it fails to recognize this distinction and consequently conflates the descriptive and normative issues.

So, Fine's argument is immune from Kukla's first criticism but what about the second concerning the development of new forms of argumentation? Here I think Kukla is on much stronger ground for there does not seem to be any conclusive way of ruling out such a possibility. To be sure Fine can argue that there is no way of resurrecting realism based on a correspondence theory of truth or instrumentalism based on an observable/unobservable distinction but does this really exhaust the possibilities vis-à-vis the realism-antirealism issue? Answering this question depends largely on what one is prepared to accept as a genuine example of realism or antirealism. For example, although Fine's arguments show that the correspondence theory of truth is unworkable, there are realists like Devitt who argue that their position requires nothing more than a minimalist theory of truth. If this move is acceptable then, pace Fine, there may still be mileage in the realism-antirealism issue. Of course, Fine is fully aware how slippery his opponent is. Indeed, this is precisely why he attempts to deal with a variety of post-NOA forms of 'piecemeal realism' (Fine 1991). However, in the absence of a

[12] Leplin (1986) and Giere (1988) advocate a claim very similar to this one.

knockdown argument that will demonstrate the impossibility of *any* attempt to settle the realism-antirealism debate, how can Fine be sure that he has killed off this issue for good? This presents us with our final problem:

**Problem 4:** Is there any way of ruling out the possibility that new forms of argumentation may eventually settle the realism-antirealism debate?

## 4. NOA and pragmatism

This concludes my discussion of the four problems of NOA. My claim is that any attempt to defend or revise NOA must provide some sort of answer to each of these problems. As we have already seen, Rorty (2003) suggests that his own brand of antirepresentationalist pragmatism has the resources to provide such answers. Before considering just what these answers might be and whether or not they constitute an acceptable defence of NOA, we must first respond to Fine's claim that pragmatism is just another form of antirealism.

### 4.1 Against the 'truthmongers'

Given that Fine is so keen to get beyond the debate over realism and antirealism, one might have thought that he would be sympathetic to the work of neo-pragmatists like Putnam and Rorty who also share this aim.[13] However, according to Fine, pragmatists both old and new are "truthmongers," another species of antirealist, who propose an acceptance theory of truth that:

> Portrays the truth of a statement *P* as amounting to the fact that a certain class of subjects would accept *P* under a certain set of circumstances. (Fine 1986b, p.138)

Fine's claim is that, with minor variations in how these circumstances are construed, we can find this definition of truth at the heart of Putnam's "internal realism," Rorty's "epistemological behaviourism", and Kuhn's paradigm-relative concept of truth. Fine views the move toward such theories as firstly based on a rejection of the realist's notion of correspondence and secondly as grounded in behaviourism.

This second claim is particularly important for Fine's purposes because it enables him to situate his rejection of the acceptance theory of truth alongside more familiar anti-behaviourist objections:

> Just about everyone recognizes that various special applications of behaviourism are wrong; for example, operationalism, or Watson-Skinnerism. So too, just about everyone has a sense of the basic error; namely, that behaviourism makes out everything it touches to be less than it is, fixing limits where none exist. Such, indeed, is the way of these antirealisms: they fix the concept truth, pinning it down to acceptance. One certainly has no more warrant for imposing this constraint on the basic concept of truth, however, than the operationalist has for imposing his constraints on more derivative concepts (like length or mass). (Fine 1986a, p.140)

This appeal to the similarity between behaviourism in other fields of inquiry and the acceptance theory of truth has strong rhetorical force but, in the absence of further argument, it is no more than guilt by association. One cannot simply assume that objections to special applications of behaviourism are easily translatable across disciplinary boundaries.[14]

However, Fine does have other arguments against so-called acceptance theories of truth apart from their alleged behaviouristic leanings. Thus, Fine (1989) makes it clear that one important argument against the acceptance theory of truth stems from the following two (infinite) sets of judgments:[15]

| | |
|---|---|
| (1.0) that P | (2.0) that P |
| (1.1) that it is true that P | (2.1) that it is (would be) accepted that P |
| (1.2) that it is true that it is true that P | (2.2) that it is (would be) accepted that it is (would be) accepted that P |
| etc. | etc. |

With respect to the first tower of judgments, Fine notes that (no matter what statement P is taken to stand for) each successive judgment in the tower is a benign paraphrase of its predecessor. For in accordance with the redundancy property of truth to judge that P is

[13] In a list of philosophers who share his general view of the realism-antirealism issue Fine explicitly mentions Rorty's 'epistemological behaviourism and Putnam's 'internal realism' (see Fine 1986a, p.113). Rorty is also keen to stress the similarities between his position and NOA (see Rorty 1991, p.49).

[14] This shows that Fine should have provided an argument first and pointed to possible analogies with behaviourism later.

[15] As Knezevich (1989) has noted, Fine's early attempts to produce such an argument are problematic in that they seem to rely on a particular kind of epistemic fallacy. Like Fine (1989), I am not so sure that Knezevich's criticisms are correct but one thing they certainly did achieve was to force Fine to make his argument with the 'truthmongers' much clearer than his original formulation.

the same as to judge it true that P. In other words, the infinite tower of judgements collapses to its first statement, namely 'that P'. Fine's problem with the acceptance theory of truth is that the same line of reasoning *cannot* be applied to the second tower of judgments because to judge P is *not* the same as judging that the relevant epistemic community accepts P. In other words, the second tower of judgments, in contrast to the first, will not collapse to its first statement but will instead contain infinitely many distinct judgments concerning acceptance.

Having established the conclusion that the second tower of judgments behaves differently to the first Fine could have simply ruled out the acceptance theory on the grounds that it does not preserve the redundancy property of truth. However, although Fine does indeed allude to the importance of this property, he pursues a somewhat different strategy that attempts to show that the acceptance theory makes truth judgments unintelligible. His argument for this claim can be presented in the following way:

1. An acceptance theory of truth holds that to judge P true is to judge P accepted; i.e., to judge (1.1) is to judge (2.1)

2. By the redundancy assumption, this is to judge it true that (2.1)

3. By the acceptance theory, this is to judge (2.1) accepted, i.e. to judge (2.2.)

4. By the redundancy assumption, this is to judge it true that (2.2)

5. By the acceptance theory, this is to judge (2.2) accepted, i.e. to judge (2.3)
   *Ad Infinitum...*

   Conclusion: To judge P true is to judge the truth of every one of the associated infinite tower: (2.1), (2.2), (2.3)...etc.

Thus, according to Fine, the acceptance theory requires that every single truth judgement be based on an infinite number of further judgments. The problem is that:

> As we try to climb the tower we get dizzy. After a few stages it seems to me that we no longer have a grasp on our judgments. Put in perfectly ordinary terms, understanding simply runs out. (Fine 1989, p.614)

In other words, acceptance theories cannot provide a conception of truth 'we can understand and strive for' and as a consequence fail to constitute an adequate theory of truth.

### 4.2 A neo-pragmatic response to Fine.

According to Fine there are only two plausible ways of responding to the infinite regress argument:

*The Revisionary Response*: Deny the contention that judging P is different from judging it accepted that P. (This would make the tower of infinite acceptance judgments collapse.)

*The Conservative Response*: Deny that identifying truth with acceptance implies that to judge P true is to judge P accepted. (This would make it unnecessary to climb the tower of acceptance judgments.)[16]

The first of these responses amounts to the claim that although judgments of truth appear to be distinct from judgments of acceptance (i.e. it makes sense to ask whether or not an accepted judgment is true) they are in fact the same. Thus it is *revisionary* in that it requires us to give up some well-entrenched intuitions about the notion of truth. The second response goes in the other direction suggesting that acceptance theories need not collapse the apparent gap between truth and acceptance because the former is not to be defined in terms of the latter. Thus it is *conservative* in the sense that it allows us to preserve those very intuitions that the first response asks us to ignore.

Perhaps unsurprisingly, Fine (1986a) finds neither of these responses to the infinite regress argument convincing. His rejection of the *revisionary* response is that, in suggesting that things are not as they seem, the revisionist truthmonger has merely followed the realist's essentialist strategy of trying to get at the way things really are. The problem with this being that, just like the realist, the revisionist truthmonger can give us no good reason why we should question our *prima facie* intuition that acceptance judgments are not the same as truth judgments. Fine's reasons for rejecting the *conservative* response are less clear but they appear to be based on the idea that an

[16] Fine quite rightly rules out the response that denies the claim that understanding runs out as we ascend the stages of the second tower.

acceptance theory of truth that does not define truth judgments in terms of acceptance is no theory of truth at all. Fine's only support for this claim seems to be Knezevich's (1989) insistence that acceptance theories must be given a strong reading (i.e. they must be read as advocating a revisionary response).[17] Now although this appears to beg the question against the acceptance theorist let us ignore this for the moment and instead focus on just how damaging Fine's remarks might be for our alleged truthmongers. Let us begin with Rorty.[18]

When it comes to discussing 'truth' Rorty (1979), like his great hero Dewey, is reticent to trade in definitive proclamations and often prefers to focus on the notion of 'warranted assertibility'. However, there are occasions when Rorty does appear to be offering something like an acceptance theory of truth:

> When extended in a certain way they let us see truth as, in James's phrase, "what it is better for us to believe," rather than as "accurate representation of reality." (Rorty 1979, p.10)

> To choose between these two approaches is to choose between truth as "what it is good for us to believe" and truth as "contact with reality." (Rorty 1979, p.176)

Passages such as these tend to support Fine's claim that Rorty defines truth in terms of acceptance because it seems that "what it is good for us to believe" is being offered as a definition or meaning analysis of "truth". Rorty, like James, would seem to be an acceptance theorist par excellence.

However, the more one reads Rorty (1979) the harder it becomes to endorse Fine's claim that he supports an acceptance theory of truth. This is particularly evident when one focuses on Rorty's comments concerning the appropriate relationship between epistemology and philosophy of language:

> Davidson distinguishes between philosophical projects which form part of "the theory of meaning properly so-called" and those motivated by "some adventitious philosophical puritanism." Roughly, Frege and Tarski pursued the first sort of project, whereas Russell and Carnap and Quine mingled pure theory of meaning with impure epistemological considerations – those which led them at various times and in various ways, to various forms of operationalism, verificationism, behaviourism, conventionalism. (Rorty 1979, p.259)

---

[17] Given that Fine disagrees with almost everything else Knezevich says it is strange that he should appeal to her in this way.

Here we can see Rorty making the Davidsonian point that providing a theory of semantic notions like meaning and truth has nothing to with 'impure epistemological considerations' but is rather an empirical study of how the various parts of our language hang together. [19] This shows that on pain of outright contradiction Rorty cannot, pace Fine, be offering a *theory* of truth (a semantical notion) in terms of "what it is good for us to believe" or "what we would all accept in the long run" (epistemological notions) because to do so would be to construct the very kind of 'impure' philosophy of language he rejects.

Given that the main target of the truthmonger argument is Putnam one would at least expect it to be applicable here. Indeed, there is plenty of evidence in Putnam (1981) that suggests he supports an acceptance theory of truth:

> 'Truth', in an internalist view, is some sort of (idealized) rational acceptability – some sort of ideal coherence of our beliefs with each other and with our experiences *as those experiences are themselves represented in our belief system* – and not correspondence with mind-independent or discourse-independent 'states of affairs'. (Putnam 1981, p.50)

Here Putnam seems to be endorsing something like a Peircean limit conception of truth. However, as we saw with Rorty, there is a danger in taking such comments out of context. In particular, one cannot assume that the 'is' in 'Truth *is* some sort of (idealized) rational acceptability' is to be taken as a statement of identity. In fact, there is good reason to think that this is precisely what Putnam wants to deny:

> Perhaps it will seem that explaining truth in terms of justification under ideal conditions is explaining a clear notion in terms of a vague one. But 'true' is *not* so clear when we move away from such stock examples as 'Snow is white.' And in any case I am not trying to give a formal *definition* of truth, but an informal elucidation of the notion. (Putnam 1981, pp.55-6)

This last sentence provides the key to understanding both Putnam's and Rorty's views on truth. Like James (1907/1981), both Rorty and Putnam believe that what truth *is* cannot be distinguished from the process by which we attain it. Thus, even though 'what it is good for us to believe' or 'idealized rational acceptability' should not be

---

[18] Because Fine only discusses Rorty (1979) and Putnam (1981) it would be unfair to refute Fine by appealing to later works. Thus, in this chapter, I have stuck to a discussion of these works only.

[19] This distinction between pure and impure philosophy of language is problematic once we realize that Rorty supports James and Dewey in thinking that the property of truth cannot be distinguished from the process of attaining it.

taken as definitions or synonyms for 'true', both Rorty and Putnam would claim that these statements tell us something about what truth *is* because they concern the role it plays in properly functioning epistemic communities.[20]

So it looks as though the 'acceptance theorist' label cannot be made to stick to either Rorty or Putnam. This shows that Fine's distinction between *revisionary* and *conservative* responses to the infinite regress argument is redundant for this dilemma can only be generated on the assumption that Putnam and Rorty *define* truth as acceptance. But, of course, if the argument outlined above is correct then they only *discuss* truth in terms of acceptance. As Kovach (1997) notes:

> The acceptance theory should not be intended as a meaning-analysis of 'truth'. The acceptance theory should rather be intended as a metaphysical-epistemic account of the property that 'truth' picks out. Thus, the argument rules out a class of implausible meaning-analyses, at best. (Kovach 1997, p.61)

Having said this both Putnam and Rorty can at least be faulted for not making themselves clearer. Both of these neo-pragmatists seem to share James's tendency to be rather careless when discussing truth. This is particularly the case with reference to their comments on behaviourism, which Fine mistakenly took to show that they were advocating something like an operationalist definition of truth when in fact they intended no such thing.[21]

### 4.3 A new name for some old ways of arguing

The subtitle of William James's *Pragmatism* labelled this fledgling philosophical movement as "a new name for some old ways of thinking." In a similar vein we might say that Fine's argument against the truthmongers is a new name for some old ways of arguing. Anyone familiar with James's account of truth and the opprobrium it elicited will most likely feel a sense of *déjà vu* when reading Fine. For example, there are an extraordinary number of similarities between this argument and Russell's critique of James's pragmatic theory of truth. Like Fine, Russell (1946) accuses pragmatists of illegitimately conflating metaphysical (e.g. truth) and epistemological issues (e.g. being good to believe). He argues that the pragmatic theory of truth is unacceptable because:

---

[20] As James noted, the main problem with the correspondence theory is not that it is wrong but rather that it is empty; it tells us nothing about the process of attaining true beliefs (see James 1907/1981).

[21] For Rorty's rejection of behaviourism see Rorty 1979, p. 97, p.174, and p.176.

> You must hold that your estimate of the consequences of a belief, both ethical and factual, is true, for if it is false your argument for the truth of your belief is mistaken. But to say that your belief as to consequences is true is, according to James, to say that *it* has good consequences, and this in turn is only true if it has good consequences, and so on *ad infinitum*. Obviously this won't do. (Russell 1946, p.844)

So, Russell rejects the pragmatic theory of truth for precisely the same reason that Fine rejects its modern-day counterpart, namely because an epistemic definition of this property fails to provide an account of truth we can understand and strive for. According to this common complaint, pragmatism confuses the order of being with the order of knowing.

However, as Smith (1978) and Putnam (1995) have pointed out, Russell made the now common error of taking Jamesian pragmatism as offering a *theory* of truth when he intended no such thing. In particular, contrary to its popular image, pragmatism, at least in the forms defended by James and Dewey, does not define truth in terms of what we want to believe or find comforting.[22] Dewey (1957) was quick to appreciate this possible misreading of his views on truth:

> Part of the reason why it has been found so obnoxious is doubtless its novelty and defects in its statement. Too often, for example, when truth has been thought of as satisfaction, it has been thought of as merely emotional satisfaction, a private comfort, a meeting of a purely personal need. But the satisfaction means a satisfaction of the needs and conditions of the problem out of which the idea, the purpose and method of action, arises. It includes public and objective conditions. It is not to be manipulated by whim or personal idiosyncrasy. (Dewey 1957, p.157)

As Dewey points out, 'defects in its statement' are at least partly responsible for the many misreadings and ill-informed debates that have centred on pragmatist accounts of truth. As Rorty (1991) has argued, these defects are particularly evident in James's discussions of truth.[23] It is ironic then that precisely the same kind of defects and misreading drive much of Fine's truthmongering argument.

---

[22] When discussing pragmatism and truth it is better not to separate off the views of Peirce. Although James and Dewey do not have exactly the same views on truth they are similar enough to be discussed together. The same is not true of Peirce who, in offering a limit conception of truth, rejects many of the views shared by James and Dewey.

[23] See Rorty 1991, pp.126-129.

## 5. Conclusion

James's *bette noir*, George Santayana (1905/1954), defended the value of history in terms of the possible consequences of ignoring it, "those who cannot remember the past are condemned to repeat it," he famously warned us. Given the discussion of the last section it would seem that this maxim applies as much to philosophers as it does to politicians and generals. However, despite the many similarities between the arguments of Fine and Russell, there is one very important respect in which they differ. Russell's arguments against the classical American Pragmatists reflected a fundamental disagreement about the form philosophy should take.[24] In contrast, Fine's argument with Putnam and Rorty reflects no such fundamental disagreement. As Rorty's enthusiastic endorsement of his views would suggest, many of Fine's stated aims are actually the same as those of neo-pragmatists like Putnam and Rorty (e.g. the desire to get beyond the realism-antirealism issue, etc.).

Russell could never have drawn on the work of James or Dewey to defend his work against the criticisms of his brilliant young student Wittgenstein. Given their shared aims the same is clearly not the case with Fine and the neo-pragmatists. So, if neither Rorty nor Putnam is a truthmonger then it might be possible to use their work to show how we can respond to the four problems of NOA. In particular, we might be able to produce a version of NOA that is a genuine alternative to realist and antirealist accounts of science. The next chapter explores the possibility of achieving this aim from the perspective of Rorty's antirepresentationalism.

[24] Roughly speaking, this was a clash between post-Fregean (sense-data) Empiricism and post-Darwinian (holistic) Naturalism.

# Chapter 7

# Rorty's Antirepresentationalism: NOA's Ark or Neurath's Boat?

## 1. Introduction

Most critics of NOA have concluded that Fine fails in his attempt to outline a genuine alternative to realism and antirealism. However, not everyone is prepared to give up on NOA so easily. In a recent article, Richard Rorty (2003) suggests that it must be seen as part of a larger project in philosophy, one that eschews the idea that knowledge is about how linguistic representations relate to reality:

> Fine has become famous for his defence of a thesis whose discussion seems to me central to contemporary philosophy - namely, that we should be neither realists nor antirealists. On this point he agrees with my favourite philosophers of language, Donald Davidson and Robert Brandom. I see the increasing consensus on this thesis as marking a breakthrough into a new philosophical world. In this new world, we shall no longer think of either thought or language as containing representations of reality. (Rorty 2003, p.1)

Given Rorty's notorious claim that "epistemology is dead" (Rorty 1979) it is hardly surprising that he should see Fine's polemics against realism as grist for his antirepresentationalist mill. He is particularly impressed with Fine's claim that "realism involves a profound leap of faith, not at all dissimilar from the faith that animates deep religious convictions" (Fine 1986a, p.116). Rorty takes this analogy to show that Fine is advocating a position very similar to his own according to which we must learn to do without non-human authorities like God or the independent world of scientific realism. Instead, Rorty wants us to "believe that scientific, like moral, progress is a matter of finding ever more effective ways to enrich human life" (Rorty 2003, p.2).

There are a number of problems with Rorty's attempt to enlist Fine for his antirepresentationalist project. As Rorty freely admits, there are "passages, or lines of thought, in Fine's work, which are obstacles to my syncretic efforts" (Rorty 2003, p.4). For example, Rorty draws attention to what we have been calling the problem of collapse, a problem that is a result of the claim that NOA allows us to understand "truth in the usual referential way" (Fine 1986a, p.130). As we have seen, the problem with this claim is that it is hard to see how NOA can differ from realism if it is committed to this correspondence-style view of truth. The aim of this chapter is to see whether or not Rorty can overcome this and the three other problems identified in the last chapter in order to produce a version of NOA that answers Fine's critics. This proposed makeover of NOA raises two important questions. Firstly, is it possible to offer a Rortian solution to the four problems of NOA? Secondly, even if available, are Rortian solutions to these problems consonant with Fine's overall project?

In order to answer these questions we first need to appreciate Rorty's overall philosophical project. In the next section I provide an introduction to Rorty's pragmatic philosophy in terms of three central themes: antirepresentationalism, ethnocentrism and liberalism. Here I try to show that Rorty's antirepresentationalism should not be confused with various forms of antirealism, e.g. instrumentalism, constructivism, idealism, etc. Further, I support Rorty's claim that just as antirepresentationalism should not be seen as just another form of antirealism, neither should it be seen as form of relativism. In section 3, I show how Rorty uses this antirepresentationalist account to argue for a transformation in the way we view science and its relation to the rest of culture. In section 4, I go on to show how an antirepresentationalist view of science and philosophy can be used to tackle the four main problems facing NOA. In section 5, I suggest that although Rorty can offer ways of responding to each of the four problems there is a problem concerning his views on science and its relation to the rest of culture. In particular, I argue that Rorty oscillates between a romantic (or "poeticised") view of science that has much in common with social constructivism and a pragmatic (or Davidsonian) view of science as a way of coping with rather than representing the world. I conclude by suggesting that Rorty would be better off sticking to the pragmatic account of science as prediction and control outlined in section 3. However, in section 6, I argue that even if we restrict ourselves to this account of science it still turns out to be incompatible with certain key components of Fine's NOA. In section 7, I conclude that although Rorty can provide answers to the four problems of NOA, he cannot provide an account of science that is consonant with the overall aims of Fine's project.

## 2. Antirepresentationalism, Ethnocentrism and Liberalism

Richard Rorty is arguably the most famous, some might say notorious, proponent of the view that pragmatism was, and indeed continues to be, a fruitful way of doing philosophy.[1] Thanks to Rorty and other supporters of this view, the works of Dewey, James and Peirce (and to a lesser extent Mead, Schiller and C.I. Lewis) are now rightfully recognized as significant contributions to the Western philosophical tradition.[2] No longer is it possible for Europeans to dismiss pragmatism as a naïve philosophical apology for the needs of a developing industrial nation.[3] Indeed, Rorty is keen to stress that there are common themes that link the pragmatism of James and Dewey to the work of great thinkers in the European philosophical tradition, on both sides of the analytical-continental divide. As Rorty (1982) says, "James and Dewey were not only waiting at the end of the dialectical road which analytical philosophy traveled, but are waiting at the end of the road which, for example, Foucault and Deleuze are currently traveling" (Rorty 1982, p.xviii).

One consequence of Rorty's eclectic approach to philosophy is to make his work peculiarly interesting to read. In contrast to the narrow focus of much analytical philosophy, his work teems with references to issues and debates from a large variety of disciplines ranging from ancient and early modern philosophy to existentialist theology and modern literary criticism. However, the breadth and ambition of Rorty's synoptic vision is not without its drawbacks. In the present context the most important of which is the difficulty of presenting an easily digestible summary of Rorty's philosophical output. In order to get round this problem I have chosen to follow Rorty's own suggestion that his work can be understood in terms of a commitment to three interrelated philosophical positions: antirepresentationalism, ethnocentrism and liberalism. Because our interest is mainly in the viability of a Rortian reformulation of Fine's NOA more attention will be paid to the epistemological and metaphysical aspects of this project as opposed to its overtly political and social dimensions. However, as I now hope to show, Rorty's views on the former cannot be properly understood without appreciating the role of the latter.

In *Philosophy and the Mirror of Nature*, Rorty (1979) presents us with a devastating historical critique of the Western philosophical tradition from Plato to Kant.

---

[1] Horton (2001, p.15) describes Rorty as "everybody's favourite whipping boy."

[2] It should be noted that Rorty thinks that there has been a tendency to over praise Peirce at the expense of James and Dewey. Rorty claims that this is largely due to Peirce's anticipation of certain problems central to logical empiricism – the dominant school of thought in mid-twentieth century analytical philosophy. See Rorty 1982, p.160 for more on this issue.

[3] In some sense this charge would actually have appealed to Dewey. As Rorty argues, the influence of Hegel on Dewey persuaded him that all philosophical systems were just a reflection of a particular point in world history (see Dewey 1957, p.v)

Here Rorty follows Dewey, the later Wittgenstein and Heidegger in arguing that Western philosophy is the product of a certain sort of picture, a picture that relies on maintaining distinctions between subject and object, internal and external, mind and body, fact and value, etc. Roughly speaking, the difference between 'therapeutic' philosophers like Rorty and analytic philosophers like Searle or Nagel is in how they react to Wittgenstein's claim that we are held captive by this picture. Therapeutic philosophers see this picture and the distinctions that support it as optional, one of many ways in which we might try to do philosophy. In this way, they follow Hegel and Nietzsche in seeing particular philosophical systems as being intimately connected to their historical milieu:

> Everything the philosopher has declared about man is, at bottom no more than a testimony as to the man of a *very limited* period of time. Lack of historical sense is the family failing of all philosophers. (Nietzsche 1977, p.29)

Rorty suggests that by developing a proper historical sense we can rid ourselves of the idea that we can somehow transcend the needs and interests of our own particular historical situation. For Rorty, the value of acknowledging this point is that it allows us to question the fruitfulness or necessity of discussing problems that were generated in a quite different historical setting to our own.

Analytical philosophers like Nagel will happily grant Wittgenstein's point that the Western philosophical tradition is held captive by a certain kind of picture. However, in contrast to Rorty, Nagel has a much more pessimistic view of our ability to put this picture behind us:

> Even if philosophical problems were mere manifestations of our particular historical situation or the accidental forms of our language, we probably wouldn't be able to free ourselves from them. (Nagel 1986, p.11)

Here we can see Nagel defending a much more Kantian view of philosophy according to which there are questions we know we shouldn't ask but somehow cannot help asking. Kant's solution to this problem is, in a sense, as therapeutic as that of Nietzsche and Wittgenstein. He suggests that by making ourselves fully aware of our tendency to ask unanswerable questions we may at least be able to reduce the number of occasions where we attempt to do just that. However, it is important to note that for Kant, just as for Nagel, there is no question of throwing out the picture or way of looking at the world that generates such questions. Kantian philosophical therapy, in contrast to its Wittgensteinian alternative, is a therapy for symptoms not causes. The problems handed down to us by great philosophers of the past are inescapable features of human thinking; they are constitutive of the philosophical enterprise itself. If one ceases to tackle these problems then one ceases to be a philosopher.[4]

In one sense, Rorty is likely to agree that therapeutic thinkers like Wittgenstein, Heidegger and Dewey are not really "philosophers," especially when this is used as an honorific term for those committed to the Cartesian-Kantian tradition. For Rorty, "Philosophy" in this sense is what Wittgenstein stopped doing when he rejected the picture theory of language developed in the *Tractatus*, what Heidegger stopped doing when he gave up his quest for the meaning of "being", and what Dewey gave up on when he swapped Darwin for Hegel. For Rorty, the difference between "Philosophy" and the therapeutic conception primarily lies in the way they treat the concept of *representation.* According to Rorty, the one thing that links together the work of philosophers as different as Plato, Kant and the early Wittgenstein is an attempt to work out how our concepts, knowledge or language correspond to or, more importantly, *represent* the world. In contrast, he suggests that the therapeutic philosophy of the later Wittgenstein, Heidegger and Dewey constitutes an *antirepresentationalist* account of human beings and the world around them:

> By antirepresentationalist account I mean one which does not view knowledge as a matter of getting reality right, but rather as a matter of acquiring habits of action for coping with reality. (Rorty 1991, p.1)

For Rorty, this conception of knowledge as acquiring habits of action rather than accurate representations is what links the work of the later Wittgenstein and Heidegger to the pragmatism of Peirce, James and Dewey.[5] It is nothing less than the rejection of the Greek distinction between knowing and doing; a distinction Dewey thought was generated by the need to defend established custom from the experimental knowledge of the lowly artisans in ancient Greece.[6]

---

[4] This is why many analytical philosophers do not consider figures like Nietzsche, Foucault and Derrida to be 'real philosophers'. Such figures are variously described as doing literary philosophy (Russell on Nietzsche), sociology or charlatanism.

[5] Alexander Bain, a contemporary of Peirce's at Harvard, was the first to put forward this idea of beliefs as habits of action.

[6] For Dewey's discussion of just how we arrived at this distinction see Dewey 1957, ch.1.

The claim that beliefs are ways of coping with the world rather than ways of representing the world can make antirepresentationalism sound like instrumentalism. This impression is strengthened when one recalls that Dewey used 'instrumentalism' as a label for his own brand of experimental philosophy. However, as Rorty has repeatedly stressed, antirepresentationalism must not be confused with traditional antirealist positions. As we have seen, the instrumentalist differs from his realist opponent in that he wants to distinguish between two classes of statements; one set of statements about observables and another set about unobservables. Instrumentalists then claim that only the former set of statements is representational, the latter set being useful as instruments for organizing our experience but not a representation of the way the world really is. In contrast, the antirepresentationalist does not think that *any* statement represents anything else. As Rorty puts it:

> The representationalism-vs.-antirepresentationalism issue is distinct from the realism-vs.-antirealism one, because the latter issue arises only for representationalists. (Rorty 1991, p.2)

So, the antirepresentationalist differs from both the realist and the instrumentalist because he rejects the idea that it is useful to argue about which statements accurately represent the world as opposed to which are merely instrumentally useful. If Rorty is correct, the realism-antirealism issue will only seem worthwhile if one is committed to the representationalist idea that true statements must correspond to non-linguistic states of affairs.

Rorty's desire to get rid of representationalism has also been mistaken for a form of idealism.[7] The dispute between the realist and the idealist typically centres on how we should conceive the relationship between thought or language and the world. The realist thinks that it is a world independent of human cognition that determines thought or language whereas the idealist thinks that thought or language determines, or perhaps just is, the world. Antirepresentationalism can appear to be idealist precisely because it denies the possibility of drawing meaningful distinctions between thought or language and the world. However, Rorty is again at pains to stress that antirepresentationalism should not be confused with either idealism or realism:

---

[7] One such philosopher is Bernard Williams, see Rorty 1991, p.4.

> Antirepresentationalists need to insist that "determinacy" is not what is in question – that neither does thought determine reality nor, in the sense intended by the realist, does reality determine thought. More precisely, it is no truer that "atoms are what they are because we use 'atom' as we do" than that "we use 'atom' as we do because atoms are as they are." *Both* of these claims, the antirepresentationalist says, are entirely empty. Both are pseudo-explanations. (Rorty 1991, p.5)

We are back to the idea that traditional philosophical disputes are a result of being held captive by a certain sort of picture. In this case, it is the idea that we must have some view about whether thought determines world or world determines thought. Just as with the realism-instrumentalism issue, Rorty wants to deny this common premise that makes the realism-idealism debate seem worthwhile. For better or worse, Rorty denies that "it is explanatory useful to pick and choose among the contents of our minds or our language and say that this or that item "corresponds to" or "represents" the environment in a way that some other item does not" (Rorty 1991, p.5). He simply refuses to play what he sees as a rather pointless game.

Even if antirepresentationalism is distinct from traditional positions like instrumentalism and idealism, most analytical philosophers are unlikely to be impressed with Rorty's attempt to consign large parts of the Western philosophical tradition to the dustbin of history. Given that the work of contemporary analytical philosophers largely gets its meaning from its place within this tradition this is hardly surprising, but there is more than just professional interest at work here. There is a genuine concern about any position that would simply have us turn our back on most of what our philosophical ancestors had to say about the world and our place within it. Whatever one thinks about the nature of philosophy one must have very good reason for making such a move. Otherwise one may suspect that Rorty's rejection of traditional philosophical problems is nothing more than a failure of nerve or perhaps just a result of his 'sour attitude' to the great systems of the past.[8] What reasons does Rorty give for turning our back on these problems?

Rorty thinks there are two main reasons that we should become antirepresentationalists: the historical failure of representationalist philosophy and the lack of a God's-eye standpoint. Rorty points out that hundreds and, in some cases, thousands of years of philosophical effort have failed to resolve the 'big questions' of Western philosophy, e.g. How do universals get their meaning? How does thought arise

---

[8] In advocating a 'genetic method' whereby philosophical problems are tied to their historical context, Dewey was often accused of taking a 'sour attitude' to the great systems of the past. See Dewey 1957.

out of matter? Is reality determined by thought or thought determined by reality? What is the metaphysical status of value? In Rorty's view, to say that we should no longer try to provide representationalist answers to these questions is not to take a 'sour attitude' to the work of philosophers who thought this was possible and desirable. Rather, it is to say that representationalism has had its innings but failed to live up to its billing – it is time to give something else a try. Whether or not one finds this historical diagnosis convincing is again likely to depend on general views about the nature of philosophy and what it is supposed to do. If, like Rorty, one thinks that representationalism is just one of many ways of thinking about the world then one is likely to regard our failure to resolve traditional philosophical problems as an indication that it is time for a change. If, like Nagel or Williams, one thinks that representationalism is the only way we have of thinking about the world then one is likely to see this "failure" as nothing more than an indication that these are difficult problems. Who said philosophy was supposed to be easy?

At this point, Rorty wants to say that there is a big difference between a difficult problem and impossible one. Difficult problems are those that we have some hope of solving in the future; impossible problems are those that we have no chance of solving at any time. Rorty claims that there is a very good reason why we should see the 'pseudo-problems' of representationalism as an example of the latter rather than the former. This has to do with the lack of what Putnam (1981) has called the "God's-eye point of view":

> The representationalists' attempt to explain the success of astrophysics and the failure of astrology is, Putnam thinks, bound to be merely an empty compliment unless we can attain what he calls a God's-eye standpoint – one which has somehow broken out of our language and our beliefs and tested them against something known without their aid. But we have no idea what it would be like to be at that standpoint. (Rorty 1991, p.6)

The claim here is that the fact that we speak a language and have a particular set of beliefs rules out the possibility of finding a neutral standpoint to compare them with a world that is somehow independent of our cognition of it – "the human serpent is over all" (James 1907/1981).

One oft mentioned problem with this claim is that it seems to imply a form of relativism. For if there is no point from which we can compare and evaluate different conceptual schemes or belief systems then it seems we will have to say that modern science is no better than reading tea-leaves as a way of dealing with the world, that somehow both are "true" when relativized to a particular conceptual scheme. However, as Rorty sees it, this conclusion can only be arrived at if one conflates various senses of "relativism":

> Three different views are commonly referred to by this name. The first is the view that every belief is as good as every other. The second is the view that "true" is an equivocal term, having as many meanings as there are procedures of justification. The third is the view that there is nothing to be said about either truth or justification apart from descriptions of the familiar procedures of justification which a given society - *ours* - uses in one or another area of inquiry. The pragmatist holds the ethnocentric third view. But he does not hold the self-refuting first view, nor the eccentric second view. (Rorty 1991, p.23)

So, if Rorty is correct, the lack of a God's-eye standpoint does not imply that every belief is as good as every other. Instead, it means that we can only judge other conceptual schemes or systems of belief by comparing them with our own. This "ethnocentric" view is the one behind Davidson's warning that there is no possibility of translating a language radically different from our own unless we assume that its speakers share something like our view of the world - the so-called principle of charity. It is also presumably what Wittgenstein meant when he claimed that if Lion's could speak we wouldn't be able to understand them.

Although he advertises it as one of the three senses in which "relativism" is generally used, Rorty is not happy using it as a label for the ethnocentrism that results from antirepresentationalism. Again, Rorty suggests that it is representationalism that is responsible for this confusion:

> The realist thinks that the whole point of philosophical thought is to detach oneself from any particular community and look down at it from a more universal standpoint. When he hears the pragmatist repudiating the desire for such a standpoint he cannot quite believe it. He thinks that everyone, deep down inside, *must* want such detachment, and sees him as an ironic, sneering aesthete who refuses to take the choice between communities seriously, a mere "relativist." But the pragmatist, dominated by the desire for solidarity, can only be criticized for taking his community *too* seriously. He can only be criticized for ethnocentrism, not for relativism. (Rorty 1991, p.30)

To many this distinction between ethnocentrism and relativism is likely to offer little comfort. After all, the real worry about relativism is that it seems to leave us with no way of justifying our preference for one conceptual scheme over another. In telling us that we cannot help prejudicing our own scheme but that no justification of this preference is possible ethnocentrism seems to do little more than rub salt in the wound.

However, Rorty's point is that one will only think of ethnocentrism in this way if one continues to think that we need transcendent justification for our beliefs and practices. If we give up on this idea by developing a desire for solidarity rather objectivity then we will be able to develop new ways of dealing with ethnocentrism. In particular, we will be able to see that the best way of avoiding the disadvantages of ethnocentrism lies in our being more inclusive and open to unfamiliar ideas. In other words, being more *liberal* in our dealings with other cultures. This political solution to the dangers of ethnocentrism is what links Rorty's antirepresentationalism to his political liberalism:

> Our best chance for transcending our acculturation is be brought up in a culture which prides itself on *not* being monolithic – on its tolerance for a plurality of subcultures and its willingness to listen to listen to neighboring cultures. This is the connection Dewey saw between antirepresentationalism and democracy. (Rorty 1991, p.14)

Rorty's hope is that we can reverse the traditional relationship between philosophy and our democratic habits by rejecting the idea that the latter are in need of justification by the former. Instead, Rorty suggests that we learn to find ways of coping with ethnocentrism rather than trying to transcend it, that we should replace objectivity with solidarity as the goal of our community. As we will now see, this claim has important consequences for Rorty's views on the relationship of science to the rest of culture.

## 3. Rorty on Science: Objectivity vs. Solidarity

For Rorty, the objectivist or representationalist legacy in Western thought is nowhere more obvious than in the way the realist views modern science:

> Worries about "cognitive status" and "objectivity" are characteristic of a secularized culture in which the scientist replaces the priest. The scientist is now seen as the person who keeps humanity in touch with something beyond itself. As the universe was depersonalized, beauty (and, in time, even moral goodness) came to be thought of as "subjective." So truth is now thought of as the only point at which human beings are responsible to something non-human. A commitment to "rationality" and to "method" is thought to be a recognition of this responsibility. The scientist becomes a moral exemplar, one who selflessly expresses himself again and again to the hardness of fact. (Rorty 1991, p.35)

According to Rorty, the problem with viewing science in this way is not only that it relies on a representationalist picture of how scientists relate to a nonhuman world but also that it leads us to devalue other academic disciplines. In particular, it leads us to think that disciplines that cannot produce the sort of prediction and control offered by the natural sciences must be immature or perhaps in some way fall short of the scientific ideal. For Rorty, it is this sort of positivistic thinking that has led to unfortunate attempts by literary theorists, philosophers and historians to make their discipline "more scientific."

Criticism of the positivist attempt to base the humanities on the methods of the natural sciences has its source in counter-Enlightenment Romanticism and the collapse of Hegelianism. As West (1996) points out:

> Hegel's grand system was, for all its dialectical sophistication and comprehensive synthesis, the product of an exclusively theoretical or intellectual consciousness. In response several currents of thought placed value on 'life' or existence', as opposed to the theoretical rationality demonstrated no less by Hegel than by more mainstream Enlightenment thinkers. Romantics defended spontaneity, emotion and individuality against the reductive categories of theoretical reason. Poetry rather than science provides the most adequate path to truth, the most adequate understanding of humanity. (West 1996, p.81)

This emphasis on spontaneity over method, emotion over reason and poetry over science is at the heart of the romantic reaction to the scientism of the Enlightenment. However, this resistance to the "reductive categories of theoretical reason" comes in both weak and strong forms. In its weak form, this simply amounts to the claim that the methods of the natural sciences are unsuitable for the study of humanity, life and existence but are more than adequate for the investigation of the nonhuman world. In its strong sense, it amounts to the claim that these methods are unsuitable for both sorts of investigation. We can call these two variants *weak romanticism* and *strong romanticism* respectively.

Historically speaking, weak romanticism is a product of the excesses of its stronger relation. Despite some notable exceptions, the idea of scientific inquiry based on romantic ideals failed to upset the hegemony of reason and method in science. As West (1996) notes, one reason for this was the tendency of romanticism to overstate the role of will and passion resulting in a "potentially dangerous irrationalism."[9] The first

[9] See West 1996, p.82.

serious attempt to offer a way of combining romanticism with a respect for science founded on method was put forward by Wilhelm Dilthey. Dilthey suggested that there were two contrasting ways of obtaining knowledge – explanation *(Erklaren)* and understanding *(Verstehen)*. Explanation, the goal of the natural sciences, involves subsuming phenomena under general laws - this allows us to predict and control the behaviour of objects in the world. Understanding, the goal of the human sciences, involves the recovery of meanings associated with actions, statements and human artifacts - this allows us to describe the particularity and complexity of individual lives. For Dilthey, it is as illegitimate to make explanation the goal of the human sciences, as it is to make understanding the goal of the natural sciences. The different subject matters of the natural sciences and the humanities require that they be carried out in quite different ways. In other words, Dilthey advocates a form of *weak romanticism* that acknowledges the importance of method and reason in the natural sciences.

The importance of Dilthey's distinction is that it seems to offer a middle ground between the extremes of Enlightenment scientism and strong Romanticism, a way of slipping between the horns of Lockean atomism and Goethean holism. Given his suspicion of scientism in the humanities one might have thought that Dilthey's ingenious distinction might have appealed to Rorty but this is not what we find. According to Rorty, both scientism and Dilthey's weak version of romanticism fail to provide us with a satisfactory account of how we should view the relationship between the natural sciences and the humanities:

> Neither sort of rhetoric is very satisfactory. No matter how much humanists talk about "objective values," the phrase always sounds vaguely confused. It gives with one hand what it takes back with the other. The distinction between the objective and the subjective was designed to parallel that between fact and value, so an objective value sounds as vaguely mythological as a winged horse. Talk about the humanists' special skill at critical reflection fares no better. Nobody really believes that philosophers or literary critics are better at critical thinking or at taking big broad views of things, than theoretical physicists or microbiologists. So society tends to ignore both these kinds of rhetoric. It treats the humanities as on a par with the arts, and thinks of both as providing pleasure rather than truth. (Rorty 1991, p.36)

This rejection of scientism and Diltheyan romanticism seems to leave Rorty with only one live option, namely the Goethean or strong form of romanticism whereby poetry and subjective truth replace the methods of natural science as the ultimate source of knowledge about ourselves *and* the world. Can this really be what Rorty is after?

Although Rorty's views on this matter are not always clear, his stated aim at least is to distinguish his position from this strong form of romanticism. In fact, it should come as no surprise that Rorty wants to take up a position that in some sense goes beyond or puts behind us the distinctions that made the positivism vs. romanticism debate seem worthwhile in the first place:

> These distinctions between hard facts and soft values, truth and pleasure, and objectivity and subjectivity are awkward and clumsy instruments. They are not suited to dividing up culture; they create more difficulties than they resolve. It would be best to find another vocabulary, to start afresh. (Rorty 1991, p.36)

As Rorty sees it, there is no longer any point in carrying on the debate about whether it is poetry or physics that tells us more about the world, this is just a bad question to ask. Instead, we should find ways of discussing the merits of poetry and physics that carry no suggestion of one or the other being more in touch with the way things really are. For Rorty, it is only the removal of the representationalist view of science that will allow us to transform our views in this way:

> We first have to find a new way of describing the natural sciences. It is not a question of debunking or downgrading the natural scientist, but simply of ceasing to see him as a priest. We need to stop thinking of science as the place where the human mind confronts the world, and of the scientist as exhibiting proper humility in the face of superhuman forces. We need a way of explaining why scientists are, and deserve to be, moral exemplars which does not depend on a distinction between objective fact and something softer, squishier, and more dubious. (Rorty 1991, p.36)

So, in Rorty's brave new world, any differences between the humanities and the sciences will have to be explained without the usual representationalist framework. What does this antirepresentationalist account of science look like?

Rorty begins by distinguishing between two senses of the term "rationality." In the first sense, Rorty suggests that we are prepared to call something rational if it follows procedures and methods with criteria for their success laid down in advance. Relying on a strangely Popperian view of scientific method, Rorty argues that "knowing in advance what would count as disconfirming his hypothesis and [being] prepared to abandon that hypothesis as a result of the unfavourable outcome of the single experiment" makes the scientist the paradigm of rationality, at least in the methodical sense. Rorty contrasts this methodical sense of rational with a second sense we associate with this term:

> In this sense, the word means something like "sane" or "reasonable" rather than "methodical." It names a set of moral virtues: tolerance, respect for the opinions of those around one, willingness to listen, reliance on persuasion rather than force. These are the virtues which members of a civilized society must possess if the society is to endure.... On this construction, to be rational is simply to discuss any topic – religious, literary, or scientific – in a way which eschews dogmatism, defensiveness, and righteous indignation. (Rorty 1991, p.37)

For Rorty, it is in this second sense that both science and the humanities deserve to be called rational. The problem is that if we rely on the first sense instead then only science can properly be called rational. As Rorty points out, being concerned with the specification of ends rather than means, there is no way for the humanities to specify in advance the criteria for their success.

Which sense of rationality should we use to characterize science and the humanities? Unsurprisingly, Rorty suggests that it is the second, weaker sense of rationality, that of being reasonable and tolerant rather than methodical, that should be used to assess *both* science and the humanities. In fact, Rorty wants us to give up on the first sense of being rational altogether. Rorty realizes there are powerful intuitions against the move to the weaker sense of rationality but he urges us to overcome them:

> Both humanists and the public hanker after rationality in the first, stronger sense of the term: a sense which is associated with objective truth, correspondence to reality, and method, and criteria. We should not try to satisfy this hankering, but rather to eradicate it. No matter what one's opinion of the secularization of culture it was a mistake to try to make the natural scientist into a new sort of priest, a link between the human and the nonhuman. (Rorty 1991, p.37)

Although he is keen to stress the work of Quine and Wittgenstein in freeing us from the grip of method, Rorty's views on science rely mainly on his reading of Kuhn (1970). This is largely because he thinks that phrases like "there is no theory-independent way to reconstruct phrases like 'really there'" show that Kuhn was the first to openly advocate an antirepresentationalist account of science. Indeed, in many cases, Rorty is prepared to take Kuhn as being the last word on all matters scientific. There are problems both with Rorty's over reliance on Kuhn as scientific oracle and his reading of Kuhn's account of science. However, I will save these problems for later discussion and return to Rorty's antirepresentationalist account of science. If we follow Rorty in abandoning the methodical sense of rationality as a way of describing the differences between sciences and humanities, what is left to say about these two sorts of inquiry? Are there any differences between them at all apart from subject matter?

Because Rorty rejects the idea that science has a special method or privileged relation to the world, he also rejects the idea that science is a natural kind. In other words, he doesn't think there is anything *essential* about science that marks it off from the rest of culture. However, this is not to say that Rorty rules out the possibility of drawing a purely pragmatic distinction between science and nonscience. In fact, Rorty thinks that there is a very familiar account of what science is supposed to be that will do just this:

> On the (familiar, if Whiggish) interpretation of Bacon common to Macaulay and Dewey, Baconians will call a cultural achievement "science" only if they can trace some technological advance, some increase in our ability to predict and control, back to that development. This pragmatic view that science is whatever gives us this particular sort of power will be welcome if one has developed doubts about traditional philosophical inquiries into scientific method and into the relation of science to reality. (Rorty 1991, p.47)

According to this 'Baconian-Deweyan' view of science, whether or not a particular sort of inquiry counts as science turns out to be an empirical question concerning "some increase in our ability to predict and control." Given that there is no prima facie reason why inquiries in the humanities should not do this, it is perfectly possible in Rorty's view for such achievements to count as science. Of course, in most cases this is unlikely but this is to miss Rorty's point about the nature of the science-nonscience distinction.

The point is precisely that we should stop thinking of the science-nonscience distinction has having a particular nature, that it "somehow cuts culture at a philosophically significant joint."[10] The Baconian-Deweyan account of science as offering prediction and control is not, in Rorty's view at least, a definition of what science *is*. Rather it is way of drawing a distinction between one sort of inquiry and the rest of culture in terms of what it allows us to do. However, this distinction has no epistemic or metaphysical significance, it is a purely pragmatic or sociological account of the way in which we use the term "science" to mark out certain sorts of activity. For Rorty, one could just as easily say that the difference between science and nonscience is that in the former we find it easier to reach consensus on important topics. Of course, this is not say that nonscience is less rational than science, that sociology is less worthy

[10] See Rorty 1991, p.47.

than physics. Again, Rorty wants to insist that being rational is about being tolerant or reasonable rather than the ability to point to a special method or privileged relation to reality that physics has but sociology does not. To return to the discussion of the previous section, rationality is about solidarity rather than objectivity. Or rather that when understood in Rorty's preferred terms, solidarity or intersubjectivity is as objective as we could ever, or should ever, want to be. For Rorty, this applies as much to science as it does to any other area of culture.

## 4. Antirepresentationalism and the Four Problems of NOA

I move now to the questions that motivated our discussion of Rorty's antirepresentationalist account of philosophy, science and culture. At the beginning of this chapter I suggested that any proposed makeover of NOA must be judged in terms of whether or not it can provide adequate responses to the four problems discussed in the previous chapter: the problems of acceptance, collapse, asymmetry and closure. Although Rorty (2003) is primarily concerned with showing how his antirepresentationalist account of truth would help Fine to avoid the problem of collapse, there are other features of Rorty's antirepresentationalism that suggest ways of dealing with the other three problems that face NOA. Indeed, if providing 'solutions' to these problems is the test of adequacy for a viable defence/reformulation of NOA then it is possible to make quite a strong case for Rorty's antirepresentationalist philosophy as the best way of achieving this aim. Of course, it may be that an antirepresentationalist version of NOA is objectionable on some other grounds, e.g. it clashes with the basic presuppositions or aims of Fine's project, but I will leave this issue to following sections. So, for the moment, let us concentrate what we can say about the four problems of NOA from a Rortian perspective.

### 4.1 The Problem of Acceptance

The first problem discussed in the previous chapter had to do with Fine's claim that the 'core position' of the realism-antirealism debate is a shared belief that our attitude toward the truth-status of scientific claims must not substantially differ from our attitude toward the truth-status of more mundane, everyday claims. The 'problem of acceptance' arises because this 'homely line' cannot be acceptable to many antirealists given that it asserts precisely what they want to deny, i.e. that the truth-status of scientific claims referring to unobservables does substantially differ from that of everyday claims referring to observables. What, if anything, can we say about this problem from Rorty's antirepresentationalist perspective?

The first thing to say is that despite Papineau's (1996) best efforts there really does not seem to be any way of making sense of Fine's claim that all antirealists can accept the core position. It is simply misleading to say, for example, that instrumentalists treat statements like "A quark passed through the bubble chamber" as being on a par with statements like "The cat is on the mat." So, the best we can hope for is to look for an alternative way of looking at this issue that captures the spirit, rather than the actual content of, Fine's analysis. At this point we can recall from the previous section that, like Fine, Rorty is keen to stress that there is an important connection between realism and antirealism. For Rorty, both of these positions are founded on representationalism – the idea that it is useful to contrast our thoughts, language or theories against an external world that is independent of our own making. So, in this sense at least, Rorty follows Fine in thinking that there is a 'core position' to the realism-antirealism debate. The difference being that Rorty wants to see the core position of this issue as a shared commitment to representationalism whereas Fine wants to see it as a shared belief in the homely line.

Although Fine and Rorty both attempt to single out the core of the realism-antirealism issue they do so for very different reasons. In identifying the homely line as the core position Fine attempts to cut away realist and antirealist extravagances in order to leave a nonrealist *via media* – the Natural Ontological Attitude. In contrast, Rorty wants to identify representationalism as the core position of the realism-antirealism issue in order to dismiss *both* of these positions. Indeed, from Rorty's point of view, Fine's attempt to outline an alternative to both realism and antirealism is doomed from the start. If we view both of these positions as committed to representationalism then any attempt to slip between them or provide a mediating position will necessarily fail for it will inevitably be representationalist. For Rorty, the idea of mediating or slipping between these positions is the wrong sort of metaphor to use; it is simply not radical enough. Instead, Rorty suggests that we take up an antirepresentationalist position that is defined by its opposition to a certain sort of picture that holds both the realist and antirealist captive.[11] In Rorty's view the only way to 'get beyond' or 'put this picture behind us' is to stop taking the debate between the realist and the antirealist seriously.

[11] See Rorty (1991), p.7.

So, an antirepresentationalist 'solution' to the problem of acceptance would involve two main alterations to Fine's account. Firstly, we must give up on the idea that the homely line is the core position of the realism-antirealism debate. Instead, we must think of this debate as being made possible by a shared commitment to representationalism. Secondly, and as a corollary, we must challenge Fine's view that the core position of the realism-antirealism issue is something we should keep rather than get rid of. In Rorty's view, it is only by doing this that we will put behind us the picture that holds the realist and antirealist captive – something that Fine's approach does not quite manage to do. If these changes are made to Fine's account then there is suddenly no problem with the claim that the realist and the antirealist are both committed to the 'core position.' By definition, 'representationalism' is just the name for the general philosophical presuppositions that make this debate meaningful. One can argue with Rorty's interpretation of these representationalist presuppositions or his desire to get rid of them but one surely cannot doubt that they exist.

Another way of making the difference between Fine's NOA and Rorty's antirepresentationalism clear is to say that whereas Fine wants us to accept the homely line *because it is the core position*, Rorty wants us to accept the homely line *because it is not the core position*. For Rorty, the homely line is just the idea that there is nothing much to say about the difference between talk about goldfish and talk about electrons.[12] According to this view, one cannot arrive at the homely line by trying to split the difference between realism and antirealism because this will just turn up yet another representationalist theory about how words or thoughts relate to the world. This view is lent credence by reflecting on the fact mentioned earlier that NOA has on various occasions been mistaken for both realism and antirealism. So, in this case at least, antirepresentationalism would seem to have the advantage over NOA in being easily distinguishable from its representationalist rivals. Indeed, as I will now show, this is precisely the conclusion that Rorty (2003) reaches with respect to Fine's views on truth.

## 4.2 The Problem of Collapse

Fine's approach relies on the claim that there is something at the heart of both realism and antirealism that is worth preserving, i.e. the core position. The basic idea of NOA is that once 'unnatural attachments' are stripped away from this core we will be left with a

[12] See Rorty (1991), p.52.

position everyone can endorse – a sort of commonsense realism – where truth can be understood:

> [I]n the usual referential way, so that sentence (or statement) is true just in case the entities referred to stand in the referred to relations. Thus NOA sanctions ordinary referential semantics and commits us, via truth, to the existence of the individuals, properties, relations, processes, and so forth referred to by the scientific statements that we accept as true. (Fine 1986a, p.130)

The problem of collapse arises because is hard to see how this understanding of truth and reference is supposed to differ from the realist's view of truth as correspondence with the facts. What, if anything, can we say about this problem from a Rortian perspective?

To answer this question there is no need for speculation because Rorty (2003) himself has provided a response. He suggests that we see the problem of collapse as the result of Fine's unfortunate attachment to the notion of ontological commitment:

> I suspect he drags in "ordinary referential semantics" because he thinks that the deployment of such semantics might help one to decide what ontological commitments to have. But it would accord better with the overall drift of Fine's thinking if he were to discard that unfortunate Quinean idea rather than attempting to rehabilitate it. NOA, Fine says, "tries to let science speak for itself, and it trusts in our native ability to get the message without having to rely on metaphysical or epistemological heating aids. So why, I am tempted to ask Fine, would you want to drag in a semiotic hearing aid such as ordinary referential semantics? (Rorty 2003, p.5)

So, Rorty agrees with Fine's critics that the appeal to truth "in the usual referential way" is a mistake, at least if NOA is supposed to be distinct from realism. The 'solution' to this problem (and the best way of letting science speak for itself) is to drop the notion of ontological commitment that underpins the appeal to semiotic hearing aids like "ordinary referential semantics".

Rorty suggests that the best way to do this is to follow Davidson by refusing to link up referential semantics with the notion of ontological commitment. As Rorty points out, Davidson acknowledges the importance of the concept of reference to the development of an adequate theory of truth. However, unlike Fine, Davidson does not think that using reference in this way "commits us, via truth, to the existence of the individuals, properties, relations, processes, and so forth referred to by the scientific statements that we accept as true" (Fine 1986a, p.130). For Davidson, although reference is a "posit we need to implement a theory of truth" (Davidson 1984, p.222), a theory of truth "does not explain reference, at least in this sense: it assigns no empirical content directly to relations between names or predicates and objects" (Davidson 1984, p.222). So, from a Davidsonian perspective, reference is indeed part of a theory of truth for natural languages but, pace Fine, this has no ontological significance.

As Rorty points out, there are places where Fine appears to endorse something like a Davidsonian theory of truth, for example when he says that those who accept NOA are:

> Being asked not to distinguish between kinds of truth or modes of existence or the like, but only among truths themselves in terms of centrality, degrees of belief, and the like. (Fine 1986a, p.127 quoted in Rorty 2003, p.6)

And, more explicitly:

> It will be apparent by now that a distinctive feature of NOA, one that separates it from similar views currently in the air, is NOA's stubborn refusal to amplify the concept of truth by providing a theory or analysis (or even a metaphorical picture). Rather, NOA recognizes in "truth" a concept already in use and agrees to abide by the standard rules of usage. These rules involve a Davidsonian-Tarskian referential semantics, and they support a thoroughly classical logic of inference. (Fine 1986a, p.133)

Of course, it doesn't seem to make much sense to deny that one has a theory of truth whilst endorsing Davidsonian-Tarskian referential semantics. However, perhaps we can make sense of this remark by taking Fine to be saying that if one follows Davidson one will know as much as there is to know about truth if one can derive T-sentences of the form "p is true if and only if p". In other words, when Fine says that NOA does not have a theory of truth he means to say that it does not have a *robust* account of what truth is.

So, perhaps Fine would endorse Rorty's attempt to defend NOA from the charge that it collapses into just another form of realism. Whether or not this solution does indeed "accord better with the overall drift of Fine's thinking" will be discussed later. However, for the moment we can say that adopting a Davidsonian theory of truth would at least save NOA from the charge that it is just another form of realism based on the correspondence theory of truth. Again, from a Rortian perspective, we can offer a diagnosis of why it is that Fine's attempt to outline a nonrealist view of truth collapses in this way. This is simply because Fine's analysis of the realism-antirealism debate in terms of the core position sets him off on the wrong track. No matter how much stripping away he does, Fine is unable to get rid of the representationalism that underpins the realism-antirealism issue. This point is amply demonstrated by the residual representationalism found in the claim that we should understand truth "in the usual referential way."

### 4.3 The Problem of Asymmetry

The fact that Rorty's antirepresentationalism is an attempt to put the whole realism-antirealism issue behind us rather than to find a mediating position between these two positions also suggests a quick solution to the 'problem of asymmetry'. As we saw in chapter 6, this problem is a result of Fine's Metatheorem argument – an attempt to show that instrumentalism provides a 'better explanation' than realism for the success of science. Kukla (1994) suggests that this claim is put forward on the proviso that instrumentalism will be rejected on other grounds. However, from an antirepresentationalist perspective this argument is strictly speaking unnecessary because the arguments that allow us to dismiss realism are precisely the same as those that allow us to dismiss instrumentalism:

> We pragmatists try to distinguish ourselves from instrumentalists not by arguing against their answers but against their questions. Unless one were worried about the really real, unless one had bought in on Plato's claim that degrees of certainty, or of centrality to our belief system, were correlated with different relations to reality, one would not know what was meant by "the everyday sense of existence." (Rorty 1991, p.52)

For Rorty, instrumentalism and realism are just as bad as each other, just two sides of the same representationalist coin. Rorty's treatment of them is thus symmetrical in the sense that neither is taken to provide a better explanation of the success of science than the other. On this antirepresentationalist view, Fine's attempt to privilege instrumentalism over realism is yet another indication that he hasn't gone far enough in his attempt to dismiss the realism-antirealism issue. [13]

### 4.4 The Problem of Closure

It seems then that the main problem with NOA, at least from a Rortian perspective, is that it relies on too many of the representationalist ideas that are presupposed by the realism-antirealism debate. If, as Rorty suggests, one accepts the lesson of antirepresentationalism "that neither does thought determine reality nor, in the sense intended by the realist, does reality determine thought" (Rorty 1991, p.5) one will be able to see that any attempt to find a mediating nonrealist position between realism and antirealism is doomed to failure (the problem of acceptance); the antirepresentationalist arguments that we use against realism are equally applicable to instrumentalist forms of antirealism (the problem of asymmetry); and, that the concept of reference has nothing to do with what ontological commitments one should have (the problem of collapse). However, for Rorty, escaping the grip of representationalism is neither easy nor entirely a matter of providing knockdown arguments against it. Rather, it is a matter of coming to see that "the traditional Western metaphysico-epistemological way of firming up our habits simply isn't working anymore. It simply isn't doing its job" (Rorty 1991, p.33).

This 'therapeutic' approach to philosophy suggests that we can reject the "pseudo-problems" (Rorty 1991, p.3) of representationalism on the pragmatic grounds that they ask questions it would be "culturally undesirable to exacerbate" (Rorty 1991, p.8). Rorty admits that just as he cannot conclusively rule out a resolution to the realism-antirealism debate, neither can he provide a knockdown argument that shows why antirepresentationalism is better than representationalism. However, for Rorty, this lack of a knockdown argument is entirely to be expected:

> On the view of philosophy I am offering, philosophers should not be asked for arguments against, for example, the correspondence theory of truth or the idea of the "intrinsic nature of reality." The trouble with arguments against the use of a familiar and time-honored vocabulary is that they are expected to be phrased in that very vocabulary. They are expected to show that the central terms in that vocabulary are "inconsistent in their own terms" or that they "deconstruct themselves." But that can *never* be shown. Any argument to the effect that our familiar use of a familiar term is incoherent, or empty, or confused, or vague, or "merely metaphorical" is bound to be inconclusive and question-begging. For such use is, after all, the paradigm of coherent, meaningful, literal speech. (Rorty 1989, pp.8-9)

Of course, like Fine, Rorty produces plenty of arguments against the coherence of both realism and antirealism. However, unlike Fine, Rorty clearly does not think that these arguments could ever provide conclusive reasons for rejecting either of these representationalist positions. Indeed, if Rorty is right about the necessity of phrasing

[13] Indeed, Rorty suggests that Fine and others are guilty of encouraging the instrumentalist in the belief that there is something special about the method of abduction as opposed to "the evidence of the senses." See Rorty 1991, p.52.

such arguments in terms of the very vocabulary one wants to reject, then *any* argument against representationalism is "bound to be inconclusive and question-begging."[14] On this view, the residual representationalism of NOA is entirely to be expected.

So, from a Rortian perspective, the problem of closure is nothing more than the realisation of the Kuhnian point that there are no theory-neutral criteria that will allow us to make a rational decision between two alternative paradigms or vocabularies. As Rorty says:

> Neither the realist nor her antirepresentationalist opponent will ever have anything remotedly [sic] like a knock-down argument, any more than Enlightenment secularism had such an argument against theists. (Rorty 2003, p.2)

However, pace Kukla (1994), the absence of a knockdown argument that will show us why antirepresentationalism is better than representationalism should not worry us. For Rorty, "such arguments are always parasitic upon, and abbreviations for, claims that a better vocabulary is available" (Rorty 1991, p.9). This means that the only way to convince representationalists to become antirepresentationalists is to make antirepresentationalism look more attractive than representationalism. This can be done either by pointing to the failures of representationalism or alternatively by advertising the advantages of the antirepresentationalist view. Of course, in many cases, this will amount to the same thing because in Rorty's view avoiding the "pseudo-problems" of representationalism is one of the main reasons to become an antirepresentationalist.

Let us try to sum up the value of Rorty's antirepresentationalism for dealing with the four problems of NOA. Firstly, Rorty's analysis of the realism-antirealism issue as founded on a shared commitment to representationalism presents a neat way of avoiding the problem of acceptance. According to this suggestion, Fine is right to say that there is a 'core position' of the realism-antirealism issue but is wrong to think that it is something we should try to preserve, at least if we are after a genuine alternative to realism and antirealism. Secondly, by appealing to Davidson, Rorty can show how we can speak of truth and reference without endorsing a correspondence theory of truth. Thirdly, Rorty's antirepresentationalism presents a persuasive argument against any attempt to privilege instrumentalism over realism as a better explanation of science.

[14] This is a charge that Rorty himself has encountered on many occasions, namely that he uses the tools and resources of analytical philosophy in order to argue against analytical philosophy itself.

Rorty's solution to the problem of asymmetry is simply to say that no such argument is necessary from an antirepresentationalist perspective. This is just to say that Fine should have simply stuck to the claim that the empiricist "avoids metaphysics only by committing, instead, the sin of epistemology" (Fine 1986a, p.147), rather than producing the question-begging Metatheorem argument. Finally, Rorty's therapeutic conception of philosophy allows us to respond to the problem of closure by saying that the lack of a knockdown argument against the realism-antirealism debate is irrelevant. For Rorty, such representationalist debates are to be dismissed on practical rather than argumentative grounds. That Fine might endorse such a response is suggested by his attempt to show "the essential role that nonrealist attitudes have played in the development of science in this century" (Fine 1986a, p.113). It is also suggested in the Rortian sounding claim that questions concerning the aims of science "call for empathetic analysis to get at the cognitive (and temperamental) sources of the question, and then a program of therapy to change all that" (Fine 1986a, p.148).

## 5. A Tale of Two Rortys

So, it seems that Rorty's antirepresentationalism can provide 'solutions' to the four problems of NOA. However, it remains to be shown that the Rortian account of science and philosophy that provides these solutions is consonant with Fine's overall project. In section 3, I presented a summary of Rorty's antirepresentationalist account of science suggesting that it aims to avoid three equally unsatisfactory representationalist positions: Positivism (or scientism), Strong Romanticism and Weak Romanticism. In that section, I suggested that Rorty attempts to outline a position that in someway denies the common assumptions that make these positions seem viable. However, as I will now argue, Rorty is not always consistent on this point and at times appears to be offering a view of science and its relation to the rest of culture that is barely distinguishable from the sort of strong romanticism that he elsewhere rejects. Only when we have properly discussed this side of Rorty's thinking on science, as opposed to the side discussed in section 3, will we be able to assess the compatibility of Rorty's views on science with those expressed in Fine's NOA.

A common complaint against Rorty is that he too often overstates his case. For example, Dennett has whimsically suggested that one can arrive at the 'truth' about some philosophical topic by looking at what Rorty has to say about it and then multiplying it by 0.637, the so-called 'Rorty factor'. Now although many of Rorty's readers will sympathize with the view that Rorty is prone to push things a little too far, Dennett's suggestion presupposes that one can identify a consistent position from which we can arrive at the 'truth' via the Rorty factor. Unfortunately, it is not always possible or even desirable "to assemble Rorty's scattered remarks into a comprehensive and coherent system" (Horton 2001, p.16). As many of his critics suggest, Rorty often makes seemingly contradictory remarks on a variety of topics, e.g. truth and justification (Thompson 2001), human nature (Soper 2001), relativism, etc. Indeed, even Rorty himself is now prepared to admit that in both his early and more recent work he has often "spoken in incautious ways" (Rorty in Festenstein and Thompson 2001, p.51) although he hopes that he is "gradually learning to be less aphoristic and less susceptible to accusations of paradox-mongering" (Rorty in Festenstein and Thompson 2001, p.51).

Given the "tensions, inconsistencies and contradictions" (Horton 2001, p.16) that plague Rorty's work it is hardly surprising to find that his views on science are often less than consistent. Despite his claim to be offering an antirepresentationalist account of science, Rorty often falls back on the sort of representationalist distinctions he claims to have rejected. For this reason alone it would be quite easy to dismiss Rorty's views of science as irretrievably flawed. However, I want to follow the lead of Horton (2001) and Thompson (2001) in urging that we should read Rorty on science "in the spirit in which he often treats others – taking up what is useful, pursuing what looks promising and rejecting or passing over what look to be his less impressive lines of thought" (Horton 2001, p.16). In the present context, I want to argue that a sympathetic reading of Rorty presents us with a choice between two alternative views of science and its relation to the rest of culture. Just as Rouse (1987) suggests that there are two Kuhns, I want to suggest that there are (at least) two Rortys.[15] Let me now elaborate on this claim.

In section 3, I argued that Rorty's antirepresentationalist account of science is based on the rejection of both the positivist view that the social sciences and humanities should try to become "more scientific" and the Diltheyan view that we should

[15] It might be objected that there are a lot more than two Rortys. In response to this objection I can only say that, like Rouse, my claim that there are two Rortys is restricted to his claims about science. Nothing I have to say bears on the question of how many Rortys there might be if one were to analyse his output as a whole. In the same way, Masterman's list of the various uses of paradigm in Kuhn's *Structure* has little to do with Rouse's 'two Kuhns' thesis.

distinguish between the natural sciences and the humanities in terms of a distinction between explanation and understanding. Further, I showed that Rorty was keen to avoid strong romanticism, a position that would replace scientific metaphors of finding for poetic metaphors of making at the heart of our self-image. Instead, Rorty suggests that we drop the distinctions that made such issues seem important in the first place. Thus he tells us to stop distinguishing between truths that are made as opposed to truths that are found, "hard" areas of culture vs. "soft" or "squishier" areas of culture, "out there" vs. "in here", object vs. subject, etc. According to Rorty, it is only by taking these distinctions seriously that results in the idea that one has to be either a positivist or a romantic. Similarly, it is only by taking this choice between positivism and romanticism seriously that we will be tempted into taking up some kind of mediating position like Dilthey. Rorty refuses to take representationalist distinctions seriously so he is neither a positivist nor a romantic of either the weak or strong variety. This at least is the official version of what Rorty takes his antirepresentationalism to consist in.

There are many reasons to think that Rorty fails to follow through on this official version of his antirepresentationalism, not least of which concern his claims about how we should view science and its relation to the rest of culture. For example, although Rorty tells us that he has no use for "unpragmatic questions" like "made or found?" (Rorty 1998, p.29), there are a number of occasions where it seems that he does:

> If we could ever become reconciled to the idea that most of reality is indifferent to our descriptions of it, and that the human self is created by the use of a vocabulary rather than being adequately or inadequately expressed in a vocabulary, then we should at last have assimilated what was true in the Romantic idea that truth is made rather than found. What is true about this claim is just that *languages* are made rather than found. (Rorty 1989, p.7)

If he has no use for the distinction between made and found why would Rorty say that truths and languages are *made* and not found? Is this just another one of those "incautious remarks" discussed earlier or does this indicate some deeper inconsistency at the heart of Rorty's account?

Of course, it would be both unfair and unrealistic to expect Rorty, one of the most prolific philosophical writers of recent times, to be perfectly consistent on such matters. However, even if we are charitable, there is reason to think that Rorty's occasional use of the made/found distinction betrays a genuine tension at the heart of his account. To see this consider the following remark concerning the role of poets in Rorty's (1989) 'liberal ironist' model of culture:

> A sense of human history as the history of successive metaphors would be to see the poet, in the generic sense of the maker of new words, the shaper of new languages, as the vanguard of the species. (Rorty 1989, p.20)

On other occasions, Rorty (1991) suggests that this 'poet as vanguard of the species' view should come to replace the 'scientist as priest' view that characterizes so much of post-Enlightenment thought:

> If we could ever be moved by the desire for solidarity, setting aside the desire for objectivity altogether, then we should think of human progress as making it possible for human beings to do more interesting things and be more interesting people, not as heading towards a place which has somehow been prepared for humanity in advance. Our self-image would employ images of making rather than finding, the images used by the Romantics to praise poets rather than the images used by Greeks to praise mathematicians. (Rorty 1991, p.28)

One is tempted to ask at this point just exactly how a philosophy that values poets over scientists and uses "images of making rather than finding" can possibly differ from the sort of strong romanticism discussed in section 3. In privileging the poet over the scientist isn't Rorty just perpetuating the sort of debates and issues that, at other times, he finds so pointless and shopworn?

This problem concerning Rorty's apparent engagement in issues he claims to have rejected has also been noticed by Charles Taylor (1980) who offers the following diagnosis:

> Old-guard Diltheyans, their shoulders hunched from years-long resistance against the encroaching pressure of positivist natural science, suddenly pitch forward on their faces as all opposition ceases to the reign of universal hermeneutics. (Taylor 1980, p.26)

The first thing to say here is that Rorty is not, and never has been, an "old-guard Diltheyan." Rorty is quite consistent in his rejection of both positivism and the Diltheyan distinction between hard and soft areas of culture.[16] Having said this Taylor is certainly correct to point out Rorty's tendency to see science as having lost its "hardness." For example, Rorty is fond of telling us that it was Kuhn (1970) who

[16] See Rorty (1982), p.195. Also see Rorty (1991), p.1 and p.96.

showed us that there is no epistemological or methodical difference between science and the rest of culture; that science is just as "fuzzy" as politics and poetry.[17]

Despite repeated protestations against the charge that he is preaching "universal softness," Rorty frequently appears to do just that.[18] For example, here is how Rorty (1991) responds to Taylor's claim that he still harbors a desire for universal hermeneutics:

> I still share something like this fancy, but it is not exactly a fancy of the reign of universal hermeneutics. It is rather a fantasy that the very idea of hermeneutics should disappear, in the way in which old general ideas do disappear when they lose polemical and contrastive force – when they begin to have universal applicability. (Rorty 1991, p.103)

This is puzzling. Rorty responds to Taylor's charge that he has simply replaced universal hermeneutics for universal positivism with the claim that he wants the idea of hermeneutics to disappear. However, *contra* his official position, hermeneutics disappears because it has "universal applicability" not because it is one half of a debate that it is no longer worth discussing. The problem here is that there is clearly a difference between the claim that we should no longer engage in a particular debate because both sides share a certain picture we should by now have put behind us and the claim that we should no longer engage in a particular debate because one of the sides has triumphed over the other.

To see this problem more clearly, compare Rorty's response to Taylor with an example he is fond of using as a way of illustrating the difference between antirepresentationalism and various forms of antirealism. On various occasions Rorty tells us that:

> [Pragmatists] are in a position analogous to that of secularists who urge that research concerning the Nature, or the Will, of God does not get us anywhere. Such secularists are not saying that God does not exist, exactly; they feel unclear about what it would mean to affirm His existence, and thus about the point of denying it. Nor do they have some special, funny, heretical view about God. They just doubt that the vocabulary of theology is one we ought to be using. (Rorty 1982, p.xiv)

---

[17] For someone who has a lot to say about science and who also prides himself on the eclectic and diverse nature of his philosophical approach, it is surprising that Rorty is prepared to take Kuhn as being the last word on all matters scientific.

[18] As is often the case with Rorty one suspects that he protests a little too much.

So, as with Rorty's reply to Taylor, the secularist hopes for a time when we will not feel the need to discuss questions like the existence of God. However, in marked contrast to that reply, this is not because theism or atheism has attained "universal applicability" but rather because the whole vocabulary concerning theism vs. atheism has fallen out of favour. It is one thing to say that nobody worries about questions concerning the existence of God because everybody believes in God and quite another to say that nobody has such worries because they are past caring. Similarly, one may stop talking about the difference between (scientific) explanation and (hermeneutical) understanding because one side or the other has attained "universal applicability" or because the distinction between two different types of inquiry has fallen into disrepute. In arguing for the former rather than the latter Rorty again seems to fall back into the very sorts of debate he elsewhere eschews.

Rorty's tendency to lapse into bad old romantic habits like the made-found distinction leads him to say some unfortunately idealistic things about science. For example, he claims that:

> The causal independence of quarks from human discourse is not a mark of reality as opposed to appearance; it is simply an unquestioned part of our talk about quarks. Anybody who doesn't know this fact about quarks is as unlikely to grasp what they are as is somebody who thinks that human rights were there before humans. We can say, with Foucault, that both human rights and homosexuality are recent social constructions, but only if we say, with Bruno Latour, that quarks are too. There is no point to saying that the former are "just" social constructions, for all the reasons that could be used to back up this claim are reasons that would apply to quarks as well. (Rorty 1998, p.8)

On his own antirepresentationalist terms, Rorty is not entitled to the claim that anything is constructed. There is just quark-talk and human rights talk. The former is more useful in physics laboratory and the latter is more useful in courts. Rorty should have nothing else to say here, he should certainly not have anything to do with the claim that both human rights and quarks are socially constructed (i.e. they are made rather than found) for this presupposes the very God's-eye point of view or skyhook he elsewhere rejects. Again, it is as if having followed Wittgenstein in rejecting the idea that there is some metavocabulary that will allow us to compare language games, Rorty cannot quite resist the temptation to say something about why the language games that concern making are in some way better than those that concern finding.

Perhaps then we should accept that Rorty is just another social constructivist, someone who wants to derive the determinacy of the world from societal forces and the like. But, of course, we know that this cannot quite be right because we have also seen that Rorty argues against the idea that it makes sense to think in terms of the distinction between 'made by us' and 'found out there'. So, again, it seems that we are faced with an interpretative problem concerning Rorty's "real" position. Is the "real" Rorty the one who insists that the entities of science are just as socially constructed as human rights or the one who insists that such claims are "entirely empty" pseudo-explanations?

It is here I think that it pays to think of Rorty's views on science (and for that matter philosophy) as being a combination of two quite different philosophical positions that correspond to two different sides of Rorty's philosophical character. The first of these positions is characterized by Rorty's desire to get beyond the debate between positivists and romantics by rejecting the distinctions that make such debates both meaningful and possible. For this Rorty, it is pointless to say that things are made rather than found or constructed rather than discovered. Similarly, on this view, just as we must reject the positivistic view of the 'scientist as priest' so too must we resist the temptation to replace her with another sort of cultural figure as the 'vanguard of the species'. The second position is produced precisely when Rorty succumbs to this very temptation. For this Rorty, the poet is more important than the scientist because he employs images of making as opposed to those of finding. On this view, science and the rest of culture are judged in terms of how well they conform to this poetic ideal. For ease of reference I propose that we refer to the first of these positions as encapsulating the views of Rorty the *pragmatist* and the second position as encapsulating the views of Rorty the *romantic*. How are we to decide between these two ways of viewing science and philosophy?

It would better accord with the overall drift of Rorty's thinking if he were to drop the troublesome vocabulary of 'making' and 'constructing'. Instead, Rorty should stick to the claim that just as 'reality' does not make our sentences true neither do our collective social practices. Similarly, Rorty would be better off if he were to stop seeing philosophy as a competition between the alternative self-images offered by the scientist and the poet. In other words, he should resist the temptation to replace the 'scientist as priest' view with the poet as 'vanguard of the species' view. Rather, he should stick to the claim that poets are good people to consult when we want to know about love and scientists are good people to consult when we want to know about DNA and quarks. To say this does not require us to go back to the Diltheyan idea that there is some fundamental methodological or epistemic difference between physics and literary criticism or biology and poetry, it merely suggests that different disciplines serve different interests and purposes. This is *all* Rorty should have to say about the relationship between science and the rest of culture.

## 6. Letting Science Speak For Itself

Suppose then that we view Rorty's claims about science as a strange mixture of the Deweyan view that science is a matter of prediction and control and the Derridean view that science is just another literary genre. Is either of these Rortian interpretations of science a suitable way of cashing out Fine's NOA? Given Fine's views on the realism-antirealism issue it seems clear that he would have nothing to do with the constructivist-sounding claim that science is just another form of poetry, a matter of revolutionary geniuses constructing new vocabularies. So, if we are to defend NOA from a Rortian perspective then it seems that we must rely on Rorty the pragmatist rather than Rorty the romantic; the side of Rorty that recommends that we think of science as a way of coping with the world rather than a way of representing it. Is this move away from representation perhaps what Fine has in mind when he asks us to be neither realists nor antirealists about science?

When Rorty is in pragmatic mood (as opposed to his pro-poetry, anti-science mood) he is apt to say things that chime nicely with Fine's views on science. For example, Rorty says that:

> The question of whether there really *are* human rights is, from the point of view I am proposing, as pointless as the question of whether there really *are* quarks. Human rights are no more or less "objective" than quarks, but this is just to say that reference to human rights is as indispensable to debates in the UN Security Council as is reference to quarks in debates in the Royal Society. (Rorty 1998, pp.7-8)

Compare this with Fine's view of what we should say about the reality of electrons:

> How are we to arrive at the judgment that, in addition to, say, having a rather small mass, electrons are objects "out there in the external world"? Certainly, we can stand off from the electron game and survey its claims, methods, predictive success, and so forth. But what stance could we take that would enable us to judge what the theory of electrons is *about*, other than agreeing that it is about electrons? It is not like matching a blueprint to a house being built, or a map route to a

> country road. For we are *in* the world, both physically and conceptually. That is, *we* are among the objects of science, and the concepts and procedures that we use to make judgments of subject matter and correct application are themselves part of that same scientific world. (Fine 1986, pp.131-132)

Here then it seems that Rorty and Fine are endorsing pretty much the same view of what we should say about the ontological status of entities and processes in science, namely the by now familiar claim that there is no viewpoint external to our current scientific worldview from which we can judgments about what *really* exists. On this view, there is no point asking whether there are *really* quarks (Rorty) or electrons (Fine) because "we are among the objects of science" (Fine). Questions concerning the reality of quarks and electrons are exhausted by reference to current scientific consensus on such matters - there is nothing more we can say.

Perhaps then we should follow Rorty's (2003) advice and view Fine's NOA as just an underdeveloped form of Rorty's more sophisticated antirepresentationalist account. Although it is certainly possible to view NOA as a sort of poor man's antirepresentationalism I want to suggest that any attempt to do so misses a crucial difference between these two projects. Indeed, I want to argue that although Fine and Rorty may appear to reach similar conclusions about the epistemic and ontological status of science, this masks a fundamental disagreement over the role that philosophy plays vis-à-vis the ongoing practice of science. To see this we can begin by recalling from the previous chapter that one of the prime motivations behind Fine's NOA is that it rejects any attempt to place interpretations on science. Thus, according to Fine, it is not the job of philosophy to place science in a particular context or to attempt to interpret it like a play.

In contrast, this sort of interpretation is precisely what Rorty tries to provide not just for science but for the whole of culture. As Putnam has said, it is as if Rorty sees himself as a sort of "doctor of the modern soul," a sort of philosophical prophet who (if we will just listen) will lead us away from our representationalist and religious hang-ups. Of course, Rorty denies that he says any such thing but as we have already seen there is often quite a gap between what Rorty says he is doing and what he actually does (remember the Rort factor). For example, Rorty claims that once we have replaced the representationalist vocabulary of traditional philosophy with his antirepresentationalist alternative we can stop "thinking of science as the place where the human mind confronts the world, and of the scientist as exhibiting proper humility in the face of superhuman forces" (Rorty 1991, p.36). Rorty thinks it is obvious that we would be better off if we were to stop thinking of science in this way. Of course, the acceptability of this claim largely depends on exactly what Rorty has in mind here. Is it science or philosophy that would be 'better off' if we were to make the move from representationalism to antirepresentationalism?

Despite his protestations to the contrary Rorty clearly thinks that we would *all* (i.e. scientists and philosophers) be better off if we gave up representationalism. However, it is precisely this point that shows why Rorty cannot provide an account of science that is consonant with the aims of Fine's NOA because instead 'letting science speak for itself' he decides to speak for it.

## 7. Conclusion

Although Rorty can 'solve' the four problems of NOA discussed in the previous chapter he fails to provide an account of science that captures Fine's minimalist intentions. As a result I think we must conclude that our search for a solution to the four problems of NOA must continue. Antirepresentationalism is not so much NOA's ark; it is more like Neurath's boat.

# Conclusion

It seems then that we have failed in our attempt to outline a naturalistic approach to the philosophy of science that is neither realist nor antirealist. The main problem with Fine's attempt to carry out this project is that he too often lapses back into the kind of realist-sounding terminology that has confused so many of his readers. Similarly, Rorty too often succumbs to a romantic/idealist view of science according to which science is all about the creative use of metaphors and the imagination. Of course, one should not deny that these are genuine and important features of science but to claim that this is all science is about is just too romantic to swallow. As we have seen, it is one of Rorty's faults (at least to the eye of the analytical philosopher) that he can never quite help himself from going just that little bit too far. It would be easy to conclude that despite their best intentions Fine and Rorty reject the realism-antirealism issue only to recreate it in a slightly different form. I am being over sanguine in thinking that a genuine alternative to realism and antirealism is a possibility we should continue to pursue?

I think not. Another way of looking at Fine and Rorty is in terms of their virtues rather their vices. For example, Fine makes the point endorsed by many naturalists that judgments about particular scientific claims should only be assessed from a local perspective, one in which we are engaged in some sense with the inquiry itself. Here I see Fine as following Wittgenstein in saying that certain kinds of questions and issues only make sense within the context of a particular language game. Fine teaches us that decisions about what does or does not exist is a decision to be made by engaged inquirers rather than disinterested philosophical observers. Similarly, Rorty brilliantly shows us that there is no mileage left in the idea that philosophy is just another name for the problem of representation. So, in a sense, Fine and Rorty are making the same point from different perspectives.

The problems with their accounts arise when they attempt to push their conclusions into each other's domain. Thus, Fine should not have attempted to make his conclusion about what we believe when we are part of science into a general philosophical account of what can be said to exist. Similarly, Rorty should have been more careful to distinguish between the rejection of abstract philosophical theses about science and local existence claims in science itself. Once we acknowledge this point we can follow philosophers like Habermas, McDowell, Brandom, and Putnam in acknowledging that the death of Representationalism (with a capital 'r') does not require us to give up the idea that science and language are attempts to represent (with a small 'r') the world. If one takes the naturalistic turn seriously then we should at least allow for the possibility that such an account may give us all the realism we need.

# References

Anderson, J. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Armstrong, D.M. (1973). *Belief, truth and knowledge.* Cambridge: Cambridge University Press.

Bishop, M.A. & Stich S. P. (1998). The flight to reference or how *not* to make progress in the philosophy of science. *Philosophy of Science*, 65, 33-49.

Blacker, D. et al. (1994). Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease. *Archives of Neurology*, 51, 1198-1204.

Bloor, D. (1976). *Knowledge and social imagery.* London: Routledge & Kegan Paul.

BonJour, L. (2001). Externalist theories of knowledge. In H. Kornblith (ed.) *Epistemology: Internalism and externalism.* Oxford: Blackwell, 10-35. (Reprinted from *Midwest Studies*, 5 (1980), 55-73.)

Boyd, R. (1973). Realism, underdetermination, and a causal theory of evidence. *Nous*, 7, 1-12.

Boyd, R. (1983). On the current status of the issue of scientific realism. *Erkenntnis*, 19, 45-90.

Boyd, R. (1989). What realism implies and what it does not. *Dialectica*, 43, 5-29.

Bradie, M. (1994). Epistemology from an evolutionary point of view. In E. Sober (ed.) *Conceptual issues in evolutionary biology.* Boston: MIT Press.

Braithwaite, R. B. (1953). *Scientific explanation.* Cambridge: Cambridge University Press.

Buchwald, J.Z. (1985). *From Maxwell to microphysics: Aspects of electromagnetic theory in the last quarter of the nineteenth century.* Chicago: University of Chicago Press.

Buchwald, J.Z. (1989). *The rise of the wave theory of light: Optical theory and experiment in the early nineteenth century.* Chicago: University of Chicago Press.

Burns, A., Howard, R. & Pettit W. (1997). *Alzheimer's disease: A medical companion.* Revised edition. Oxford: Blackwell Science.

Callebaut, W. (1993). *Taking the naturalistic turn.* Chicago: University of Chicago Press.

Cartwright, N. (1983). *How the laws of physics lie.* Oxford: Clarendon Press.

Churchland, P.M. & Hooker, C.A. (eds.) (1985). *Images of science.* Chicago: University of Chicago Press.

Collins, H.M. (1985). *Changing order: Replication and induction in scientific practice.* London: Sage.

Davidson, D. (1984). *Inquiries into truth and interpretation.* New York: Oxford University Press.

De Mey, M. (1982). *The cognitive paradigm.* Dordrecht: Reidel.

Descartes, R. (1968). *Discourse on method and the meditations* (F.E. Sutcliffe, Trans.). London: Penguin. (Original work published 1637)

Devitt, M. (1991). Aberrations of the realism debate. *Philosophical Studies*, 61, 43-63.

Dewey, J. (1957). *Reconstruction in philosophy.* Boston: Beacon Press (Original work published 1920).

Dewey, J. (1958). *Experience and nature.* 2nd edition. New York: Dover.

Dilthey, W. (1976). *W. Dilthey: Selected writings,* H.P. Rickman (ed.). Cambridge: Cambridge University Press.

Doppelt, G. (1986). Relativism and the reticulational model of scientific rationality. *Synthese*, 69, 225-252.

Doppelt, G. (1990). The naturalist conception of methodological standards in science: A critique. *Philosophy of Science*, 57, 1-19.

Festenstein, M. & Thompson, S. (eds.) (2001). *Richard Rorty: Critical dialogues.* Cambridge: Polity Press.

Fine, A. (1986a). *The shaky game: Einstein, realism, and the quantum theory.* Chicago: Chicago University Press.

Fine, A. (1986b). Unnatural attitudes: Realist and instrumentalist attachments to science. *Mind*, 95, 149-179.

Fine, A. (1989). Truthmongering: Less is true. *Canadian Journal of Philosophy*, 19 (4), 611-616.

Fine, A. (1991). Piecemeal realism. *Philosophical Studies*, 61, 79-86.

Fine A. (1996). Science made up: constructivist sociology of scientific knowledge. In P. Galison and D. Stump (eds.) *The disunity of science: Boundaries, contexts and power*. Stanford: Stanford University Press.

Fine, A. (1998). The viewpoint of no-one in particular. *Proceedings of the American Philosophical Association*, 72 (2), 9-20.

Freedman, K. (1999). Laudan's naturalistic axiology. *Philosophy of Science*, 66, 526-537.

Fuller, S. (1988). *Social epistemology.* Indianapolis: Indiana University Press.

Fuller, S. (1993). *Philosophy of science and its discontents* (2nd ed.). New York: Guildford Press.

Fuller, S. (1994). Mortgaging the farm to save the (sacred) cow. *Studies in History and Philosophy of Science*, 25 (2), 251-261.

Gettier, E. (1963). Is justified belief true knowledge? *Analysis*, 23, 121-123.

Giere, R. (1985). Philosophy of science naturalised. *Philosophy of Science*, 52, 331-356.

Giere, R. (1988). *Explaining science: A cognitive approach.* Chicago: Chicago University Press.

Goldman, A. (1986). *Epistemology and cognition.* Harvard: Harvard University Press.

Goldman, A. (2001). The internalist conception of justification. In H. Kornblith (ed.) *Epistemology: Internalism and Externalism.* Oxford: Blackwell, 36-67. (Reprinted from *Midwest Studies*, 5 (1980), 27-51).

Golinski, J. (1998). *Making natural knowledge: Constructivism and the history of science.* Cambridge: Cambridge University Press.

Hacking, I. (1983). *Representing and intervening.* Cambridge: Cambridge University Press.

Hacking, I. (1999). *The social construction of what?* Cambridge, MA: Harvard University Press.

Hardin, C. & Rosenberg, A. (1982). In defence of convergent realism. *Philosophy of Science*, 49, 604-615.

Hawthorn, J. (1988). Not a metatheorem in Fine. *Mind*, 97, 585-587.

Horton, J. (2001). Irony and commitment: An irreconcilable dualism of modernity. In M. Festenstein and S. Thompson (eds.) *Richard Rorty: Critical dialogues*. Cambridge: Polity Press.

Hull, D. (1988). *Science as a process: An evolutionary account of the social and conceptual development of science.* Chicago: Chicago University Press.

Hume, D. (1975). *Enquiries concerning human understanding and concerning the principles of morals*. L.A. Selby-Bigge and P.H. Nidditch (eds.). Oxford: Oxford University Press. (Original work published 1748)

James, W. (1981). *Pragmatism: a new name for some old ways of thinking*. Indiana: Hackett. (Original work published in 1907)

Jennings, R. (1989). Scientific quasi-realism. *Mind*, 98, 225-245.

Jorm, A. F. (1990). *The epidemiology of Alzheimer's disease and related disorders*. London: Chapman and Hall.

Kant, I. (1968). *Critique of pure reason* (Norman Kemp Smith, Trans.). London: MacMillan. (Original work published 1781)

Kenny, A. (1973). *Wittgenstein.* London: Penguin Group.

Kitcher, P. (1992). The naturalists return. *The Philosophical Review* 101, (1), 53-114.

Kitcher, P. (1993). *The advancement of science.* Oxford: Oxford University Press.

Kitcher, P. (1995). Author's response. *Philosophy and Phenomenological Research*, 55 (3), 653-673.

Kitcher, P. (1997). An argument about free inquiry. *Nous*, 31(3), 279-306.

Kitcher, P. (1998). Truth or consequences? *Proceedings of the American Philosophical Association*, 72 (2), 49-63.

Kitcher, P. (2000). On the explanatory role of correspondence truth. Retrieved September 25, 2001 from http://www.columbia.edu/~psk16/correspondence_truth.htm

Kitcher, P. (2001). *Science, truth and democracy.* Oxford: Oxford University Press.

Knezevich, L. (1989) Truthmongering: An exercise. *Canadian Journal of Philosophy*, 19 (4), 603-609.

Knorr-Cetina, K.D. (1981). *The manufacture of knowledge.* Oxford: Pergamon Press.

Kornblith, H., (ed.) (1985). *Naturalizing epistemology.* Cambridge: MIT Press.

Kornblith, H., (ed.) (2001). *Epistemology: Internalism and externalism.* Oxford: Blackwell.

Kovach, A. (1997). Why we should lose our NOAs. *Southern Journal of Philosophy*, 35 (1), 57.

Kripke, S. (1980). *Naming and necessity.* Oxford: Blackwell.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: Chicago University Press.

Kukla, A. (1994). Scientific realism, scientific practice and the NOA. *British Journal for the Philosophy of Science*, 45 (4), 955-975.

Kukla, A. (1997). Review: Beyond positivism and relativism. *British Journal of the Philosophy of Science*, 48, 447-454.

Kukla, A. (2000). *Social constructivism and the philosophy of science*. London: Routledge.

Lakatos (1978). *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.

Latour, B. (1987). *Science in action*. Milton Keynes: Open University Press.

Latour, B. & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. London: Sage.

Laudan, L. (1977). *Progress and its problems*. Berkeley: University of California Press.

Laudan, L. (1981). The pseudo-science of science. *Philosophy of the Social Sciences*, 11, 173-98.

Laudan, L. (1984). *Science and values*. Berkeley: Berkeley University Press.

Laudan, L. (1986). Some problems facing intuitionist meta-methodologies. *Synthese*, 67, 115-129.

Laudan, L. (1987). Progress or rationality? The prospects for normative naturalism. *American Philosophical Quarterly*, 24, 19-31. Reprinted in D. Papineau (ed.) *The philosophy of science*. Oxford: Oxford University Press.

Laudan, L. (1990a). Aim-less epistemology? *Studies in History and Philosophy of Science*, 21 (2), 315-322.

Laudan, L. (1990b). Normative naturalism. *Philosophy of Science*, 57, 44-59.

Laudan, L. (1990c). *Science and relativism: Some key controversies in the philosophy of science*. Chicago: Chicago University Press.

Laudan, L. (1996a). A confutation of convergent realism. In D. Papineau (ed.) *The philosophy of science.* Oxford: Oxford University Press. (Reprinted from *Philosophy of Science*, 48 (1981), 19-48.)

Laudan, L. (1996b). *Beyond positivism and relativism: Theory, method and evidence.* Boulder, Colorado: Westview Press.

Laudan, L., R. Laudan & A. Donovan (1992). *Scrutinizing science*. Baltimore: Johns Hopkins University Press.

Leplin, J. (1986). Methodological realism and scientific rationality. *Philosophy of Science*, 53, 31-51.

Leplin, J. (1990). Renormalizing epistemology. *Philosophy of Science*, 57, 20-33.

Leplin, J. (1997). *A novel defence of scientific realism.* Oxford: Oxford University Press.

Levi, I. (1995). Cognitive value and the advancement of science. *Philosophy and Phenomenological Research*, 55 (3), 619-627.

Lewis, P. (2001). Why the pessimistic induction is a fallacy. *Synthese*, 129, 371-380.

Lugg, A. (1986). An alternative to the traditional model? Laudan on disagreement and consensus in science. *Philosophy of Science*, 53, 419-424.

Lynch, M. (1985). *Art and artifact in laboratory science: A study of shop work and shop talk in a research laboratory.* London: Routledge and Kegan Paul.

Machamer, P. (1995). Kitcher and the achievement of science. *Philosophy and Phenomenological Research*, 55 (3), 629-636.

Matheson, C. (1996). Critical notice: Philip Kitcher The advancement of science. *Canadian Journal of Philosophy*, 26 (3), 463-489.

Maxwell, G. (1962). The ontological status of theoretical entities. In H. Feigl and G. Maxwell (eds.) *Scientific explanation, space and time*. Minneapolis, MN: University of Minnesota Press.

McMullin, E. (1987). Explanatory success and the truth of theory. In N. Rescher (ed.) *Scientific inquiry in philosophical perspective*. Lanham: University Press of America.

Merton, R.K. (1973). *The sociology of science.* Chicago: Chicago University Press.

Miller, R.W. (1995). The advancement of realism. *Philosophy and Phenomenological Research*, 55 (3), 637-645.

Moore, G.E. (1903). *Principia ethica.* Cambridge: Cambridge University Press.

Moore, G.E. (1995). Proof of an external world. In P.K. Moser and A. vander Nat (eds.) *Human Knowledge: Classical and contemporary* approaches, 383-395. Oxford: Oxford University Press. (Original work published in 1939)

Musgrave, A. (1996). NOA's ark – Fine for realism. In D. Papineau (ed.) *The philosophy of science*. Oxford: Oxford University Press. (Reprinted from *Philosophical Quarterly*, 39 (1989), 383-398.).

Nagel, E. (1961). *The structure of science.* New York: Harcourt, Brace & World.

Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.

Newton-Smith, W. (1978). The underdetermination of theory by data. *Proceedings of the Aristotelian Society*, 52 (supplement), 71-91.

Nietzsche, F. (1977). *A Nietzsche reader.* (Translated and selected by R. J. Hollingdale). Middlesex: Penguin.

Niiniluoto, I. (1977). Realism, relativism, and constructivism. *Synthese*, 89, 135-162.

Papineau, D. (ed.) (1996a). *The philosophy of science*. Oxford: Oxford University Press.

Papineau, D. (1996b). Introduction. In D. Papineau (ed.) *The philosophy of science*. Oxford: Oxford University Press.

Pickering, A. (1984). *Constructing quarks*. Chicago, IL: University of Chicago Press.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.

Psillos, S. (2000). The present state of the scientific realism debate. *British Journal for the Philosophy of Science*, 51 (4), 705-728.

Putnam, H. (1975). *Mathematics, matter and method,* vol.1. Cambridge: Cambridge University Press.

Putnam, H. (1978). *Meaning and the moral sciences*. London: Routledge.

Putnam, H. (1981). *Reason, truth and history.* Cambridge: Cambridge University Press.

Putnam, H. (1995). *Pragmatism: An open question*. Oxford: Blackwell.

Quine, W.v.O. (1951). Two dogmas of empiricism. In *From a logical point of view*. Cambridge, MA: Harvard University Press, 20-46.

Quine, W.v.O (1960). *Word and object.* Cambridge, MA: MIT Press.

Quine, W.v.O. (1974). *The roots of reference.* LaSalle, IL: Open Court.

Quine, W.v.O. (1985). Epistemology naturalized. In H. Kornblith (ed.) *Naturalizing epistemology.* Cambridge, MA: MIT Press/Bradford Books. (Reprinted from *Ontological relativity and other essays.* New York: Columbia University Press, 1970.)

Reisch, G.A. (1998). Pluralism, logical empiricism, and the problem of pseudoscience. *Philosophy of Science*, 65, 333-348.

Rorty, R. (1979). *Philosophy and the mirror of nature.* Princeton: Princeton University Press.

Rorty, R. (1982). *Consequences of pragmatism.* Minneapolis: University of Minnesota Press.

Rorty, R. (1989). *Contingency, irony and solidarity.* Cambridge: Cambridge University Press.

Rorty, R. (1991). *Objectivity, relativism and truth: Philosophical Papers Volume 1.* Cambridge: Cambridge University Press.

Rorty, R. (1998). *Truth and progress: Philosophical papers volume 3.* Cambridge: Cambridge University Press.

Rorty, R. (2003). A pragmatist view of contemporary analytic philosophy. Retrieved March 15, 2003 from *http://www.stanford.edu/~rrorty/pragmatistview.htm.*

Rosenberg, A. (1990). Normative naturalism and the role of philosophy. *Philosophy of Science*, 57, 34-43.

Rosenberg, A. (1996). A field guide to recent species of naturalism. *British Journal for the Philosophy of Science*, 47, 1-29.

Rouse, J. (1987). *Knowledge and power: Toward a political philosophy of science.* Ithaca: Cornell University Press.

Rouse, J. (1988). Arguing for the natural ontological attitude. In A. Fine and J. Leplin (eds.) *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1. East Lansing, MI: Philosophy of Science Association

Rouse, J. (1996). *Engaging science: How to understand its practices philosophically.* Ithaca and London: Cornell University Press.

Russell, B. (1946). *A history of western philosophy.* London: George Allen and Unwin.

Santayana, G. (1954). *The life of reason.* London: Constable. (Original work published 1905)

Shapere, D. (1995). Kitcher on advancing science. *Philosophy and Phenomenological Research*, 55 (3), 647-651.

Shapin, S. & Schaffer, S. (1985). *Leviathan and the Air-pump: Hobbes, Boyle and the experimental life*. Princeton, NJ: Princeton University Press.

Siegel, H. (1990). Laudan's normative naturalism. *Studies in the History and Philosophy of Science*, 21 (2), 295-313.

Smart, J.J.C. (1963). *Philosophy and scientific realism*. London: RKP.

Smith, J. (1978). *Purpose and thought: The meaning of pragmatism*. London: Hutchinson.

Solomon, M. (1995). Legend naturalism and scientific progress: An essay on Philip Kitcher's The advancement of science. *Studies in History and Philosophy of Science*, 26 (2), 205-218.

Solomon, M. (2001). *Social empiricism.* Cambridge, MA: MIT Press.

Soper, K. (2001). Richard Rorty: Humanist and/or anti-humanist. In M. Festenstein and S. Thompson (eds.) *Richard Rorty: Critical dialogues.* Cambridge: Polity Press.

Stempsey, W. E. (2000). *Disease and diagnosis: Value dependent realism.* Dordrecht: Kluwer.

Stroud, B. (1985). The significance of naturalized epistemology. In H. Kornblith (ed.) *Naturalizing epistemology.* Cambridge, MA: MIT Press/Bradford Books. (Reprinted from *Midwest Studies in Philosophy,* vol. VI (1981), 455-471.).

Stump, D. (1992). Naturalized philosophy of science with a plurality of metamethods. *Philosophy of Science*, 59, 456-460.

Taylor, C. (1980). Understanding in human science. *Review of Metaphysics*, 34, 25-38.

Thagard, P. (1988). *Computational philosophy of science.* Cambridge, MA: MIT Press/Bradford Books.

Thagard, P. (1992). *Conceptual revolutions.* New Jersey: Princeton University Press.

Thompson, S. (2001). Richard Rorty on truth, justification and justice. In M. Festenstein and S. Thompson (eds.) *Richard Rorty: Critical dialogues.* Cambridge: Polity Press.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon Press.

van Fraassen, B. C. (1985). Empiricism in philosophy of science. In P.M. Churchland and C.A. Hooker (eds.) *Images of science*. Chicago: Chicago University Press.

West, D. (1996). *An introduction to continental philosophy.* Oxford: Blackwell.

Williams, B. (1985). *Ethics and the limits of philosophy.* London: Fontana Press.

Williams, B. (1987). *Descartes: The project of pure enquiry.* London: Penguin.

Wittgenstein, L. (1953). *Philosophical investigations* (G.E.M. Anscombe, Trans.). G.H. von Wright and G.E.M. Anscombe (eds.). Oxford: Basil Blackwell.

Worrall, J. (1982). Scientific realism and scientific change. *The Philosophical Quarterly*, 32, 201-231.

Worrall, J. (1989). Fix it and be damned: A reply to Laudan. *British Journal of the Philosophy of Science*, 40, 376-388.

Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica* 43/1-2: 99-124.

Worrall, J. (1994). How to remain (reasonably) optimistic: Scientific realism and the "Luminiferous Ether." In D. Hull, M. Forbes and R. Burian (eds.) *PSA*, vol.1. East Lansing, MI: Philosophy of Science Association.

Wylie, Alison (1986). Arguments for scientific realism: The ascending spiral. *American Philosophical Quarterly*, 23 (3), 287-297.