ON THE RELATION BETWEEN THE NORMATIVE AND THE EMPIRICAL IN THE PHILOSOPHY OF SCIENCE

Timothy Childers
Department of Philosophy, Logic and Scientific Method
London School of Economics and Political Science

Thesis submitted for the PhD degree of the University of London
1996

ABSTRACT

The relation between the normative and the empirical in the philosophy of science is examined by investigating apriori and aposteriori approaches to methodology. The apriori is usually equated with the prescriptive, and the aposteriori with the descriptive. It is argued that this equation is mistaken, and that neither a purely apriori nor a purely aposteriori approach to methodology can succeed. Methodologies based on probability are used as illustrations.

Purely apriori and purely aposteriori approaches are examined in Parts I and II respectively. The former are investigated through the intuitionism of J.M. Keynes and the analytic method of Carnap. Dutch Book arguments are also considered as apriori arguments. I conclude that an apriori approach is irredeemably flawed, in that it can never meet the goal it sets for itself of producing a self-evidently justified set of rules for science. Purely aposteriori approaches are investigated in the second Part by focussing on R. Giere's and W.V.O. Quine's proposals for a naturalised epistemology. It is argued that a purely empirical approach is caught on the horns of a dilemma: if it is defended on aposteriori grounds then the argument is circular, and if on apriori grounds it is self-refuting. Thus it is shown that the aposteriori approach too cannot serve as the foundations for methodology.

However, I shall argue that Quine's project has been misunderstood, and that in fact Quine argues for aposteriori methodology from conventionalist grounds. The possibility of a conventionalist approach to the philosophy of science which avoids the problems of the purely empirical and of the purely apriori approach is explored in the third Part of this thesis. Karl Popper's early advocacy of such a conventionalist approach is discussed. The final chapter is devoted to showing how certain flaws in Popper's and Quine's conventionalist approaches may be mended. It is concluded that the conventionalist approach to methodology provides an adequate framework for the relation between the normative and the empirical in the philosophy of science.

INTRODUCTION

The epistemology of science is a species of epistemology: scientific knowledge is just knowledge of a particular kind. But what, exactly, is it that epistemologists should do? This thesis is an attempt to survey and assess the main answers to this question. I divide the answers into three classes: First, it has been claimed that the business of the epistemologist is to find the correct way to obtain knowledge, independently of scientific investigation. Having found the method of true learning, the business of philosophers becomes the promulgation of the results of the investigation in the form of methodologies. This view might seem related to revealed theology: the philosopher, like the priest, comes to know the truth, and then attempts to set others free with it. I shall call this the apriori conception of epistemology.

The second answer is that philosophy can never hope to outshine the glory of science, and it would befit philosophers to follow the example of their betters. Perhaps epistemology could be a kind of science itself, to be replaced, eventually, by an interdisciplinary enterprise such as cognitive science. This view could be seen as related to natural theology: the truth is found by contemplation of the world, in a joint enterprise of philosophers and scientists. I shall call this view the aposteriori conception of epistemology.

The third position is that the business of epistemology is to point out the implications of given rules of scientific conduct. The rules themselves are mere proposals, or perhaps have already been adopted by some scientists for various reasons. The philosopher plays the role of the therapist, helping scientists to meet goals they set themselves. This view can be seen as related to agnosticism, since the philosopher makes no claims to know, or even to be able to discover, the true path to knowledge. I shall call this view conventionalism.

This thesis is accordingly divided into three sections, each corresponding with one of the above views. The argument of the thesis is that the first two approaches cannot be successful, while the third can be. I attempt to support it by examining two types of apriorism (intuitionism, represented by J.M. Keynes, and the analytic method, represent by R. Carnap), two types of aposteriorism (championed by R. Giere and W.V. Quine), and argue that their failures show that neither view is tenable. I shall also argue that conventionalism is immune to the criticisms to which the apriori and the aposteriori approach fall (and also suggest that this may help explain why most proponents of both of these views have at some point smuggled in some form of conventionalism). But, I shall also argue that conventionalism too has problems. The challenge then becomes one of formulating an acceptable conventionalism, so I shall examine two proposals put forward by Quine and by K.R. Popper. The last chapter of this thesis is a tentative suggestion along lines suggested by these authors.

Such a grand project, that of showing that only conventionalism has so far survived, and that it is the only view that can survive, is obviously impossible (even neglecting the fact that it is undertaken by a single doctoral candidate). I limit my project in two ways. First, I concentrate only on what might be called empiricist epistemology by restricting the scope of the thesis to methodology. Secondly, I for the most part consider only probabilistic methodologies. I feel that much will not be overlooked given these constraints since most of what I say about probabilities applies *mutatis mutandis* about other methodologies.

Specifically: Part 1 deals with the issues around methodology considered as an apriori discipline. Chapter 1 covers intuitionism, with the main protagonists being J.M. Keynes and G.E. Moore. Chapter 2 deals with the analytic method proposed by Rudolf Carnap. I argue that both attempts are failures. In Chapter 3 I argue that the further failures in this programme show apriorism to be an unsuccessful enterprise.

Part II deals with the purely empirical approach. I examine and dismiss in Chapter 4 what may be called 'naive' empiricism (or 'naturalism,' as empirical methodologies are sometimes called). Only one person, R. Giere, really seems to have ever held this view: W.V.O. Quine, who is often cited as a proponent of this view instead holds, I shall argue in Chapter 5, like Carnap, an empirical-cum-conventionalist view. While I shall later argue that this view is broadly correct, I conclude that Quine's version is too narrow conceived.

The shortest section of this thesis is Part III, which is an attempt to show how to work towards an acceptable conventionalism. Chapter 6 is an argument against Karl Popper's apriorism-cum-conventionalism. Chapter 7 is an attempt to show that conventionalism is an acceptable alternative.

Acknowledgements

I cannot imagine that there could be a finer tutor than Peter Urbach: even if I could list his many qualities, it would be too long to include here. His detailed and careful criticisms of this thesis were invaluable, as is his thoughtful advice and steadfast support.

Colin Howson and Thomas Uebel have both discussed many of the issues in this thesis with me. I appreciate their advice and encouragement. Professor Nancy Cartwright discussed with me many issues related to those found in this thesis, and helped me obtain both financial support and work.

I wish I knew how to thank properly my wife, Hope Marie Childers, and my parents. I am grateful for their support and forbearance of the vicissitudes of academic life, and for their love.

I am grateful to the members of the Philosophy of Science Research Students Group for organising the best series of seminars I have ever attended. I am particularly grateful to Marco del Seta and Samet Bagce for their constant willingness to engage in ferocious argument in convivial surroundings.

Husain Sarkar first introduced me to the issues with which this thesis is concerned, and encouraged me to study at the LSE. Ms. Theresa Hunt and Mrs. Pat Gardner helped in negotiating necessary bureaucratic hurdles. The All-London Centre for the Philosophy of the Natural and Social Sciences provided me with a desk. Many friends in London have supported me in difficult times: Imogen Planner, Dagmar Lorenz-Meyer, Joanna Bennet and Graham Combi in particular. I thank them all.

Finally, I thank the members of the Working Group in Logic, Institute of Philosophy, Academy of Sciences of the Czech Republic: particular thanks are due Vladimír Svoboda, Ondrej Majer and Petr Kolář, for showing me my new home.

PART I: METHODOLOGY AS APRIORI

In this Part, I consider the apriori approach to methodology, and address the question of whether or not it can provide a sound basis. Advocates of this approach make a strong and appealing case. They reject the aposteriori approach on the grounds that values cannot be deduced from facts, nor prescription derived from description, and argue that how science actually proceeds implies nothing about how science should proceed. They also claim that any methodology must have an apriori basis. Some of the most forthright proponents of this apriori approach have been philosophers interested in constructing an inductive calculus. If such an inductive calculus were a species of logic, it would be the strongest candidate for an apriori account of methodology (if anything is apriori, it seems that logic is).

The following three chapters examine a variety of attempts to found an apriori inductive methodology on the probability calculus. They follow, in roughly chronological manner, attempts to address the problem of the nature of the normative and its relation to the empirical: Philosophers have long despaired of deducing oughts from is's. An associated worry is what sorts of things "oughts" actually are. "Is's" are, it seems, somewhat more familiar objects (known by empirical means, hence the term "natural"), whereas all we seem to know about "oughts" is that they are not "is's" (hence the term "non-natural"). Since it is hard to say what "oughts" are, it is similarly difficult to say how we come to know that there are "oughts", since, presumably we cannot come to know them in any empirical manner (or, it seems, they would be "is's"). One very popular answer to this puzzle is that we can perceive "oughts", although, of course, this perception is not of the familiar type: this type of perception is "intuition." Thus intuitionism is an epistemological counterpart of the ontological doctrine of non-naturalism. In the first chapter I examine the J.M. Keynes's attempt to found a probabilistic methodology on the basis of intuitions: my conclusion will be that these attempts are not successful.

The subject of Chapter 2 is the attempt to found an apriori inductive methodology by means of the analytic method, that is, to demonstrate that probability values can be determined solely from the meanings of the statements involved. This view is an application of a general view of justification developed in opposition to intuitionism. Rudolf Carnap is the best known proponent of this view, and I will examine his programme, and argue that it could not succeed.

In Chapter 3 I discuss Dutch Book arguments as apriori arguments to show that inductive reasoning should be reasoning in accordance with the probability calculus. I will argue that, despite some remarkable progress, these attempts have failed. I will further argue that these failures are indicative of a more general failing of the non-empirical approach.

CHAPTER 1: Keynes's Intuitionistic Methodology

I begin with J.M. Keynes for two reasons, one historical and one philosophical. The first is that Keynes's work opened a new chapter in methodology. Keynes's *A Treatise on Probability* was the first book-length work in English on the foundations of probability for 55 years (as Keynes himself notes in the introduction); Carnap credits Keynes with "the first attempt to construct an axiom system for probability" (Carnap 1950, p. 338);* and Keynes presents a justification of his inductive methodology which seems to be new, and which was adopted by many succeeding writers. The philosophical reason is that Keynes serves as a useful starting point because he held that the probability calculus was justified apriori, and his reasons are shared by many apriorists. I will argue that Keynes did not provide an adequate justification of this view, and thus many apriorist positions are also not justified.

In this chapter I will first examine Keynes's epistemology; in particular, in section A, I argue that Keynes held that we intuit probability relations. In section B, I discuss Keynes's version of the Principle of Indifference, upon which he founded his probabilistic methodology. This necessitates a discussion of the paradoxes to which the Principle of Indifference leads. Keynes's attempt to resolve the paradoxes is difficult to follow. I discuss two interpretations of Keynes's solution to the paradoxes offered by Gillies and by Howson and Urbach, and argue for an alternative interpretation. I shall then argue that Keynes's attempt to reformulate the Principle of Indifference so that it was not paradoxical was based on his account of intuitions, and that this attempt was not successful. The reason for this failure is Keynes's intuitionism, which, I shall argue in section C, was derived from G.E. Moore's account of ethical intuitionism. I argue that the problems encountered by Moore's intuitionism are even more serious for Keynes's. I conclude that intuitionism cannot provide an adequate justification for methodology.

*To distinguish between Carnap's 1950 Logical Foundations of Probability and his 1950 "Empiricism, Semantics and Ontology" I shall note the reprinting date in brackets for the latter (i.e. 1950[1967]). I shall follow this convention throughout.

A. Keynes on the Intuition of Probabilities

In his *A Treatise on Probability,* Keynes attempted to show that probability relations are objective, and can be known by intuition. He held that probabilistic inference was a branch of logic, and that probability relations are objective, and not psychological, just as logical relations are:

> ...in the sense important to logic, probability is not subjective. It is not, that is to say, subject to human caprice. A proposition is not probable because we think it so. When once the facts are given which determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively, and is independent of our opinion. The Theory of Probability is logical, therefore, because it is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals, which may or may not be rational. (Keynes 1921 [1973], p. 4)

The question then arises of how we gain knowledge of these putative objective relations. Keynes's answer provides the foundation of his apriori methodology.

Before I examine Keynes's proposed foundations of probability, it will be useful to make clear the sources of his account. Keynes was strongly influenced by Bertrand Russell and G.E. Moore: in this section I show that Keynes adopted parts of Russell's epistemology. (Russell indeed reports in his autobiography: "I had no contact with him [Keynes] in his political and economic work, but I was considerably concerned with his *Treatise on Probability,* many parts of which I discussed with him in detail." (Russell 1967, p. 71)) In particular, I will now argue that Keynes adopted Russell's epistemology of logic, which, I shall argue later, resembles in many respects Moore's epistemology of ethics.[1]

Russell held that there were self-evident truths, "incapable of demonstration," among which are "the principle of induction", "other logical principles" such as "the law of contradiction", and "self-evident truths... immediately derived from sensation." (Russell 1912 [1946], pp. 112-113) According to Russell, there are degrees of self-evidence, and "truths of perception and some of the principles of logic have the very highest degree of self-evidence..." (Russell 1912 [1946], p. 117), and moreover that "...the highest degree of self-evidence... is really an *infallible guarantee of truth.*" (Russell 1912 [1946], p. 118, my italics) In Russell's terminology, a proposition has the highest degree of self-evidence

---

[1]It is not necessary to explore the relations between these two accounts, since I am interested only in explicating Russell's position to clarify Keynes's: however, my criticisms of Moore's and Keynes's epistemology will also apply, *mutatis mutandis,* to Russell's.

if it is known by "acquaintance," and hence, what is known by acquaintance is infallibly true: "Thus this sort of self-evidence [acquaintance with a fact] is an absolute guarantee of truth" (Russell 1912 [1946], p. 137).

Knowledge by acquaintance is defined by Russell in the following way: "I say that I am *acquainted* with an object when I have a direct cognitive relation to that object, i.e. when I am directly aware of the object itself." (Russell 1910 [1917], p. 209) Exactly what Russell meant by "direct cognitive relation" or "directly aware" is not clear. Russell says that acquaintance "may be called perception, though it is by no means confined to objects of the senses." (Russell 1912 [1946], p. 136)) This perception seems to be a kind of intuition, since Russell refers to "intuition" and "intuitive knowledge." (Knowledge by acquaintance seems closely related to Descartes's "clear and distinct" ideas.) Russell gives "sense-data" as examples of things known by acquaintance: "When we ask what are the kinds of objects with which we are acquainted, the first and most obvious example is *sense-data* [sic]. When I see a colour or hear a noise, I have direct acquaintance with the colour or the noise..." (1910 [1917], p. 210) Russell also includes objects of introspection, such as universals, as things known by acquaintance. (ibid, p. 212) Russell contrasts knowledge by acquaintance with knowledge by description. The latter he characterises as the knowledge of an object that it is "*the* so-and-so, i.e., when we know that there is one object, and no more, having a certain property; and it will be generally be implied that we do not have knowledge of the same object by acquaintance." (ibid, pp. 214-215) Russell gives as an example of knowledge by description our knowledge of "the man with the iron mask": "We know that the man with the iron mask existed, and many propositions are known about him; but we do not know who he was." (ibid, p. 215) Presumably Russell means by this that we are certain that there was a person who was "the man with the iron mask", but we only know this through the accounts of others who report actually having seen the man.

It is clear that Keynes adopted these parts of Russell's epistemology in his epistemology of probability relations. Keynes uses the same terminology as Russell, drawing a similar distinction between direct and indirect acquaintance, and citing Russell's example of sense data to illustrate the former:

> We start from things, or various classes, with which we have, what I choose to call without reference to other uses of this term, *direct acquaintance*... The most important classes of things with which we have direct acquaintance are our own sensations, which we may be said to *experience*, the ideas or meanings, about which we have thoughts and which we may be said to *understand*, and facts or characteristics or relations of sense-data or meanings, which we may be said to *perceive*;—experience, understanding, and perception being three forms of direct

> acquaintance. (Keynes 1921 [1973], p. 12)

Keynes draws a distinction between direct acquaintance and direct knowledge. The difference between the two is that the former is the indubitable sensation of experiencing something, the latter is concerned with indubitable propositions. Necessary to this distinction is a difference between "[t]he objects of knowledge and belief" and "the objects of direct acquaintance" which are "sensations, meanings, and perceptions." (ibid, p. 12) Direct acquaintance, according to Keynes (and Russell), leads to direct knowledge "as the result of contemplating the objects of acquaintance." (ibid, p. 12) Keynes gives as an example of direct knowledge gained from direct acquaintance the transition from the sensing of yellow to knowledge of certain propositions about yellow:

> ...From acquaintance with a sensation of yellow I can pass directly to a knowledge of the proposition 'I have a sensation of yellow'. From acquaintance with a sensation of yellow and with the meanings of 'yellow', 'colour', 'existence', I may be able to pass to a direct knowledge of the propositions 'I understand the meaning of yellow', 'my sensation of yellow exists', 'yellow is a colour'. *Thus by some mental process of which it is difficult to give an account*, we are able to pass from direct acquaintance with things to a knowledge of propositions about the things of which we have sensations or understand the meaning. (ibid, p. 13, my italics)

(Russell used the sensation of yellow as a paradigmatic example of things known by direct acquaintance (Russell 1910 [1917], pp. 212-213); we will see later that Moore also used this example.) Keynes followed Russell in calling propositions obtained directly "self-evident": "For the objects of certain belief which is based on direct knowledge, as opposed to certain belief arising indirectly, there is a well-established expression; propositions, in which our rational belief is both certain and direct, are said to be *self-evident*." (Keynes 1921 [1973], p. 18) Keynes, like Russell, assumed direct knowledge ("knowledge by acquaintance") to be indubitable: "...I have assumed that all direct knowledge is certain." (ibid, p. 17) Keynes was never specific about the extent of direct knowledge: "About our own existence, our own sense data, some logical ideas, and some logical relations it is usually agreed that we have direct knowledge." (ibid, p. 14) Presumably "some logical ideas" refers to the axioms of the *Principia Mathematica*, although this is not clear. Keynes included probability relations among the logical relations of which we have direct knowledge. He held that we could come to know probability relations (which are second order, or "secondary" propositions) by "perceiving... [a] probability relation" which holds between the propositions of direct knowledge:

> Now our knowledge of propositions seems to be obtained in two ways: directly, as the result of contemplating the objects of acquaintance; and indirectly, *by*

> *argument*, through perceiving the probability-relation of the proposition, about which we seek knowledge, to other propositions. In the second case, at any rate at first, what we know is not the proposition itself but a secondary proposition [i.e., a proposition asserting a probability relation] involving it. When we know a secondary proposition involving the proposition *p* as a subject, we may be said to have indirect knowledge *about p*. (ibid, pp. 12-13)

And Keynes held that just as we sense yellow, so too do we sense and acquire direct knowledge of logical relations (in which, as already noted, Keynes included probability relations). He held that we have direct knowledge of logical relations, and it is these relations which allow us to make inferences from propositions obtained from direct knowledge:

> ...by the contemplation of propositions of which we have direct knowledge, we are able to pass indirectly to knowledge of or about other propositions... We pass from a knowledge of the proposition *a* to a knowledge about the proposition *b* by perceiving a logical relation between them. With this logical relation we have direct acquaintance. The logic of knowledge is mainly occupied with a study of the logical relations, direct acquaintance with which permits direct knowledge of the secondary proposition asserting the probability-relation, and so to indirect knowledge about, and in some cases of, the primary proposition. (ibid, p. 13)

So for Keynes, direct acquaintance serves two purposes: it supplies propositions about which we are certain, and it supplies knowledge of the logical and probabilistic relations we can use to reason from the certain propositions to probable propositions.

Keynes also adopted parts of G.E. Moore's epistemology, as will be documented in section C of this chapter. In particular, Keynes held that the probability relation was not definable for the reason that it could not be analysed "in terms of simpler ideas" (This is clearly inspired by G.E. Moore's intuitionist view that "good" is indefinable, as I will show later):

> [§]8. A *definition* of probability is not possible, unless it contents us to define degrees of the probability-relation by reference to degrees of rational belief. We cannot analyse the probability relation in terms of simpler ideas. As soon as we have passed from the logic of implication and the categories of truth and falsehood to the logic of probability and the categories of knowledge, ignorance, and rational belief, we are paying attention to a new logical relation in which, although it is logical, we were not previously interested, and which cannot be explained or defined in terms of our previous notions. (ibid, p. 8)

Keynes argued for this position on the grounds that, at the time of his writing, no satisfactory definition of probability had been found. While Keynes admitted that his view that the probability relation is indefinable was "incapable of positive proof," he nevertheless felt that both the difficulties involved in finding such a definition, and what he claimed to be the familiarity of the notion of probability, weighed in favour of such a view: "The presumption in its favour must arise partly out of our failure to find a definition, and partly because the notion presents itself to the mind as something new and independent..." (ibid, p. 8) (Keynes seems to have been unaware of the representation of probability as a fair betting quotient, and dismissed the classical definition of probability.) However, Keynes weakened his position somewhat from arguing that a definition of probability is indefinable to arguing (a few sentences later) that one might be discovered "in the end." Keynes thought, however, that no great harm would result from proceeding without a definition of probability, since he believed that the probability relation was easily perceived:

> Even if a definition is discoverable in the end, there is no harm in postponing it until our enquiry into the object of definition is far advanced. In the case of 'probability' the object before the mind is so familiar that the danger of mis-describing its qualities through lack of a definition is less than if it were a highly abstract entity far removed from the normal channels of thought. (ibid, pp. 8-9)

I shall criticise Keynes's intuitionist account of the probability relation as an indefinable, intuitively perceivable logical relation in section C. However, it is necessary to discuss exactly which probability relations Keynes claimed to perceive.

B. Keynes on the Principle of Indifference

As we have seen, Keynes claimed that probability relations were a kind of logical relation, and in fact, he thought of probability as a kind of partial deduction. He attempted to make the probability calculus practically useful by using the Principle of Indifference to assign numerical values to at least some probability relations. It was well known at the time that Keynes was writing that the Principle of Indifference leads to contradictions, and so he attempted to formulate a version of the Principle which avoided these difficulties. I will first discuss Keynes's solution, and then argue that his solution is unsuccessful.

Keynes believed that, in general, probabilities are not comparable: "I maintain... that there are some pairs of probabilities between the members of which *no* comparison of magnitude is possible..." There are, however, Keynes held, special cases in which probability differences can be more or less exactly determined: "...we can say, nevertheless, of some pairs of relations of probability that the one is greater and the other less, although it is not possible to measure the difference between them; and that in a very special type of case, to be dealt with later, a meaning can be given to a *numerical* comparison of magnitude." (ibid, pp. 36-7) The "very special type of case" to which Keynes refers is that in which the Principle of Indifference can be applied.

The Principle of Indifference, originally formulated by Laplace, states that in situations where we have n mutually exclusive and exhaustive hypotheses $h_1,\ldots,h_n$, and no reason for believing that any one of the alternatives is more likely than any other, we should assign equal probability to each of them. In the discrete case, if the members of the set $\{h_i\}_{i\in n}$ are exhaustive and the h_i are mutually exclusive, then by the probability calculus, $\Sigma_i \mathrm{p}(h_i)=1$, and hence the Principle of Indifference gives $\mathrm{p}(h_i) = 1/n$. For example, in the discrete case if you are contemplating the outcome of a roll of a six-sided die, which you believe to be evenly weighted, and if only one of six possible outcomes of a throw of the die can occur (i.e., excluding the possibility of the die landing on an edge or a corner, or not landing at all), the Principle of Indifference implies that the probability of a throw yielding, say, a five, is 1/6.

The Principle of Indifference can be extended to the continuous case. Consider a real-valued parameter T which ranges over the interval $[a,b]$, and suppose that in light of our present knowledge we have no reason to believe that T will take one value rather than another. Then according to the continuous form of the Principle of Indifference, the probability that T lies in the subinterval $[c,d]$, $\mathrm{p}(c\leq T\leq d)$, is equal to $\dfrac{|d-c|}{|b-a|}$. An illustration of this is a wheel of fortune for which we have no reason for believing that the pointer will stop at one point position rather than another. (This wheel is idealised, since the pointer may halt at any point on the circumference, and we believe it to be, say, uniformly lubricated, balanced, and so forth). Suppose we wish to determine the probability that the pointer will fall in a region which is one quarter of the circumference of the wheel. Denote the length of the circumference of the wheel as c. The Principle of Indifference states that the probability of the pointer falling in the area marked 0 to $\frac{c}{4}$ is $\dfrac{|\frac{c}{4}-0|}{|c-0|} = \dfrac{1}{4}$. (This is illustrated in figure 1.)

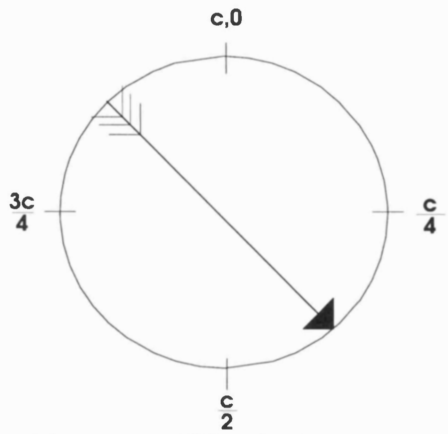

Figure 1 - Wheel of Fortune

1. The paradoxes

It was well known at the time of Keynes's writing that these forms of the Principle of Indifference generate inconsistencies. Keynes was well aware of these, and indeed made an extensive survey of the various examples of inconsistencies generated by the Principle (Keynes 1921 [1973], p. 46, also see Gillies 1973, p. 11 and 1988, pp. 186-187 for an exposition): One of the simpler examples is the bookmark paradox, which shows that inconsistency of the discrete form of the Principle of Indifference: if a book is taken at random from a library which only has, say, three colours of bookmarks (red, green and blue, with exactly one bookmark per book), then, since we have no more reason to believe that the bookmark will be red than not, the Principle of Indifference implies that the probability of selecting a red bookmark is one half. But by similar reasoning the probability of getting a blue bookmark is also one half, as is the probability of getting a green one. But this violates the third axiom of the probability calculus: the probability of mutually exclusive and exhaustive alternatives must sum to one.

Bertrand's paradox is one of the most famous examples of inconsistencies generated by the continuous form of the Principle. (Expositions may be found in von Mises 1957 [1981], p. 77, Gillies 1973, pp. 11-12 and Székely 1986, pp. 43-48). von Mises attributes the name 'Bertrand's paradox' to Poincaré, while Keynes (1921 [1973], p. 48-49) discusses a similar paradox which he attributes to von Kries. Suppose that you are interested in the relative volumes of water and wine in a particular mixture, and that you are sure that the mixture contains at least three times as much of one liquid as the other. So according to your information, the ratio of wine to water lies between three parts of wine to one part of water, and three parts of water to one part of wine. Consider the probability that the ratio of wine to water is less than or equal to 2. Our knowledge does not extend beyond $1/3 \leq \text{wine/water} \leq 3$ and so we have no reason for thinking that the ratio is one value rather than another. Hence applying the continuous form of the Principle of Indifference, $p(\text{wine/water} \leq 2) = \dfrac{|2 - \frac{1}{3}|}{|3 - \frac{1}{3}|} = \dfrac{5}{8}$. But consider now the reciprocal ratio of water to wine.

The probability that it is greater than or equal to 1/2, by the application of the Principle of Indifference, yields $p(\text{water/wine} \geq 1/2) = \dfrac{3 - \frac{1}{2}}{\frac{1}{3}} = \dfrac{15}{16}$. But these two results contradict one another, since wine/water ≥ 1/2 is the same state of affairs as 2 ≥ water/wine and hence should have the same probability: the Principle of Indifference has yielded different probabilities for logically equivalent propositions in violation of the probability calculus. (The similar example discussed by Keynes (1921 [1973], pp. 48-49) concerns the specific gravity and its reciprocal, the specific volume, of a particular substance.)

Another set of examples of inconsistencies generated by the Principle of Indifference are the "paradoxes of geometric probability." Keynes (1921 [1973], p. 51) attributes the following example of such a paradox to Bertrand, and Neyman (1952, p. 15) refers to "the so-called Bertrand's problem." I follow Borel's exposition (1950 [1965], p. 87). The problem is to determine the probability that a chord drawn at random onto a circle of radius r, will be shorter than the side of an equilateral triangle inscribed in the circle.

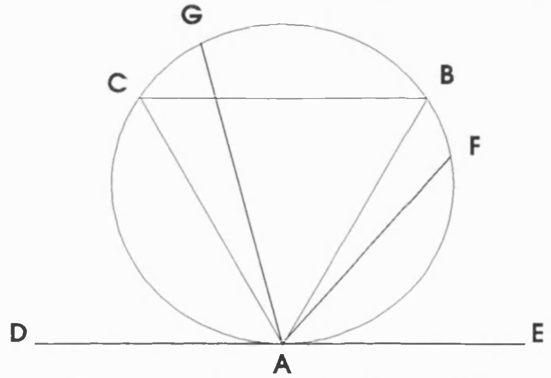

Figure 2a Geometric Probability

There are at least three ways of calculating the probability using the Principle of Indifference, illustrated in figures 2a, b and c. In the first, we might label the point at which one end of the chord falls A, and apply the Principle of Indifference to the position where the other point falls (F or G in the above figure). Consider the line DE tangent to the circle at the fixed endpoint A. This and the chord will form an angle ∠FAD (or ∠GAD). If this angle is between 60 and 120 degrees, the chord AG will be longer than the side of an inscribed equilateral triangle ABC, otherwise it will be shorter (AF). So, applying the Principle of Indifference, the chance that the chord will be shorter than the side of an inscribed triangle is 60/180 = 2/3.

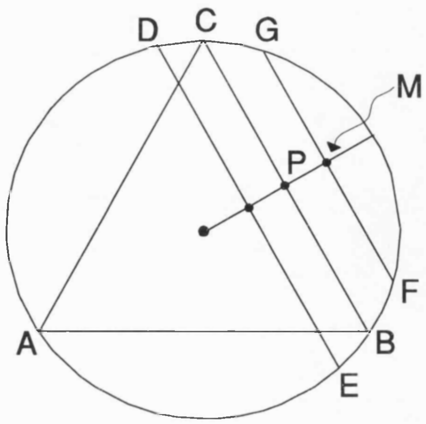

Figure 2b Geometric Probability

In the second case (figure 2b), the Principle of Indifference can be applied to where the midpoint of the chord falls with respect to a fixed direction. The chord must fall in *some* direction. Consider an inscribed triangle ABC with side BC parallel to this direction. We can determine the length of the chord as follows: a line drawn perpendicular to BC from the centre of the circle to the circumference of the circle will have its midpoint where it intersects the triangle (P). If the midpoint of the chord (M, for the chord FG) falls between the midpoint of the perpendicular line and the circle, the length of the cord will be less than that of the side of the triangle (as is FG). Similarly, if the midpoint falls less than half the distance, it will be longer (as is DE). So the probability of the chord's length being less than the side of an inscribed equilateral triangle is one half.

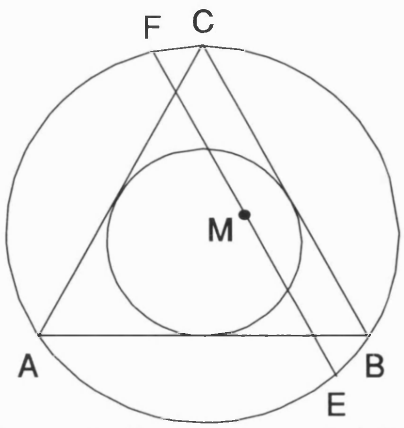

Figure 2c Geometric Probability

A third way of calculating the probability of the length of the chord being shorter than the side of an inscribed equilateral triangle is shown in Figure 2c. Considers a circle inscribed in the inscribed triangle, and denote the midpoint of the chord as M. Since the inscribed circle has radius (1/2)r, if the midpoint falls that circle, the chord (FE in the figure) will be longer than the side of an inscribed triangle. The area of the smaller circle G is 1/4 of the area of the larger circle, and so the probability that the chord will be smaller is 3/4. Hence the Principle of Indifference assigns at least three different probabilities to one proposition, and, once again, generates a contradiction.

Bertrand's paradox and the wine/water paradox show that the Principle of Indifference is not invariant with respect to non-linear transformations such as inversion. Consider any real-valued parameter θ about which we know only that it may take any value in the interval $[a,b]$. The Principle of Indifference generates the probability density distribution $\int_a^b \frac{1}{b-a} d\theta$. Now consider f = θ^2, so that $a^2 \leq f \leq b^2$. From this we obtain $d\theta$ = df/2θ.

Substituting df for $d\theta$ gives us $\int_{a^2}^{b^2} \frac{1}{2\theta(b-a)} d\phi$. But since $\theta = \sqrt{\phi}$, we have

$\int_{a^2}^{b^2} \frac{1}{2\sqrt{\phi}(b-a)} d\phi$. However, according to the Principle of Indifference, given a parameter

f which ranges from a^2 to b^2, we should have the distribution $\int_{a^2}^{b^2} \frac{1}{(b^2-a^2)} d\phi$. So the

distributions generated by the Principle of Indifference are not invariant under transformation of the variables to which it is applied. As Keynes pointed out, these paradoxes arise since, in general, $\frac{|d-c|}{|b-a|}$ is not the same as $\frac{|f(d)-f(c)|}{|f(b)-f(a)|}$:

> In general, if x and $f(x)$ are both continuous variables, varying always in the same or in the opposite sense, and x must lie between a and b, then the probability that x lies between c and d, where $a<c<d<b$, seems to be $\frac{(d-c)}{(b-a)}$, and the probability that $f(x)$ lies between $f(c)$ and $f(d)$ to be $\frac{[f(d)-f(c)]}{[f(b)-f(a)]}$. These expressions, which represent the probabilities of necessarily concordant conclusions, are not, as they ought to be, equal. (Keynes, 1921 [1973], p. 51)

It is clear that Keynes was aware both of the paradoxes and their causes. The next section details his attempts to circumvent them.

2. Keynes's solution to the paradoxes

It was vital for Keynes's project of providing a sound apriori basis for a probabilistic methodology that he be able to assign probabilities unambiguously to at least some propositions, but it was clear to him that the only method available for this was badly flawed. Therefore, Keynes needed to reformulate the Principle to avoid the paradoxes. But this is by is by no means easy, and, unfortunately, Keynes's reformulation is far from clear. In this section, I shall present two interpretations of what Keynes meant (offered by Gillies, and by Howson and Urbach). I shall argue that neither is entirely satisfactory, and propose an alternative interpretation. However, none of the interpretations suggest that Keynes had succeeding in formulating a Principle of Indifference free from paradoxes.

At the centre of Keynes's approach was his theory of intuition:

> [§]12. Without compromising the objective character of relations of probability, we must nevertheless admit that there is little likelihood of our discovering a method of recognising particular probabilities, without any assistance whatever from intuition or direct judgment. Inasmuch as it is always assumed that we can sometimes judge directly that a conclusion *follows from* a premiss, it is no great extension of this assumption to suppose that we can sometimes recognise that a conclusion *partially follows from*, or stands in a relation of probability to, a premiss. Moreover, the failure to explain or define 'probability' in terms of other logical notions, creates a presumption that particular relations of probability must be, in the first instance, directly recognised as such, and cannot be evolved by rule out of data which themselves contain no statements of probability. (ibid, pp. 56-57)

To discuss Keynes's version of the Principle of Indifference we must first consider his notion of "preference" and "relevance." "Preference" is not intended in the usual sense relating to the desirability of an object, but is a probabilistic notion. Using modern notation, x is preferred to y on evidence e if $p(x|e) > p(y|e)$, that is, x is more likely than y given e. e_1 is, in Keynes's terminology, irrelevant to h on evidence e if there is no proposition e_x such that $e \wedge e_1 \vdash e_x$, and e $\nvdash e_x$, and $p(h|e) \neq p(h|e \wedge e_x)$. That is, a statement is irrelevant to a hypothesis, given some evidence, if neither the statement nor anything derivable from the statement in conjunction with the evidence, ~~does not~~ affects the conditional probability of the hypothesis. Clearly, relevance is a more fine-grained notion than preference. Using this terminology, we can restate the condition that the Principle of Indifference be applied only in cases where there is no reason for believing one alternative hypothesis more likely than another: we should be indifferent between hypotheses only if there is no evidence which is relevant to one hypothesis, but not the other; that is, we should be indifferent between two hypotheses h_1 and h_2 only if there is no evidence relevant to h_1 which is not also relevant to h_2. However, this is only a necessary condition for the applicability of the principle. Keynes states a second condition, that the hypotheses must be non-decomposable and "of the same form" (I will discuss this shortly). These two conditions are jointly necessary and sufficient for the application of the Principle of Indifference.

Keynes relied on intuitions to determine whether or not there is some evidence which is relevant to a hypothesis: for each hypothesis under consideration it is necessary to make a "direct judgment" as to whether relevant information exists:

> The principle [of indifference] states that 'there must be no known reason for preferring one of a set of alternatives to another'...
> Before... then, we can begin to apply the Principle of Indifference, *we must have made a number of direct judgments* to the effect that the probabilities under consideration are unaffected by the inclusion in the evidence of certain particular details. We have no right to say of any known difference between the two alternatives that it is 'no reason' for preferring one of them, unless we have judged that a knowledge of this difference is irrelevant to the probability in question. (ibid, p. 58, italics mine)

Hence, intuitions, so-called "direct judgements", were the foundation for Keynes's application of the Principle of Indifference, and hence for his probabilistic methodology.

But the paradoxes of the Principle are not generated by mistaken judgements of irrelevance. Rather, judgements of irrelevance determine those cases in which the Principle might be applied. The paradoxes arise from changes of description. Keynes aimed to avoid this by further restricting the application of the Principle of Indifference to those cases in which the hypotheses we are indifferent between are not further decomposable, or, in Keynes's terminology, "divisible." Keynes attempted to specify what decompositions he had in mind: a sentence is decomposable if it is equivalent to a sentence which is a disjunction of statements which are (1) exhaustive, (2) mutually exclusive and (3) have positive probability. (ibid, p. 65) Using modern notation, we can express these three criteria as

> A sentence h is *decomposable* if there exist sentences $h_1,\ldots,h_n$ such that
> (i) $\mathrm{p}(h \equiv (h_1 \vee \ldots \vee h_n) | e) = 1$,
> (ii) $\mathrm{p}(h_i \wedge h_j | e) = 0$ for all $i, j \leq n$
> (iii) $\mathrm{p}(h_i | e) \neq 0$, for all $i \leq n$.

Keynes acknowledged that statements can always be decomposed in a trivial way that seems to threaten his solution: statements of the form R are equivalent to $((R \wedge A) \vee (R \wedge \neg A))$, where A is any statement. This satisfies the three conditions above, and so it would seem that Keynes's solution will not work. However, Keynes points out that the propositional functions are not of the same form. (While Keynes does not actually define sameness of form, it seems a relatively simple matter to do so: all one would need to do would be to define a normal form, and then give a recursive means of transforming sentences into that normal form, and an algorithm for comparing the sentences thus transformed. As I shall show in chapter 2, this is how Carnap defined decomposibility.)

A criticism can be levelled against Keynes's condition of non-decomposability, viz, that it is language-dependent, since decomposability depends on which constants we have in our language, and so is dependent on the descriptive richness of our language. In the first case we only have two colours, and in the second we have three colours. Thus, unless Keynes believes that there are ultimate constituents of the world which we can intuit, our decompositions must always be relative to our current set of beliefs about the properties from which the world is composed. I will argue below that this is a serious shortcoming of Keynes's theory.

It seems that Keynes's reformulated Principle of Indifference is directed to paradoxes such as the bookmark paradox where the possibility space is discrete. If it were not for the problem of language dependence, then, it seems that Keynes did rehabilitate, to some extent, the Principle of Indifference (although I will later criticise Keynes's solution in the discrete case). However, Keynes was unsuccessful in avoiding paradoxes in the continuous case. Keynes noted that in the continuous case, however finely the possibility space is partitioned, the remaining alternatives can always be further decomposed. Hence, the Principle of Indifference is not applicable in these cases:

> [§]24. It is worth while to add that the qualification of §21 [i.e., the three preconditions above] is fatal to the practical utility of the principle of indifference in those cases only in which it is possible to find *no* ultimate alternatives which satisfy the conditions... It is often the case, however, that we cannot by any process of finite sub-division arrive at indivisible sub-alternatives, or that, if we can, they are not on the evidence indifferent. In the examples given above, for instance,... where x is a part of unspecified magnitude in a continuum, there are *no* indivisible sub-alternatives. (1921 [1973], p. 67)

But Keynes seems to have believed that the paradoxes in the continuous case could be overcome with his modified Principle of Indifference, although he is not at all easy to understand on this point. His claim evidently is that the paradoxes in the continuous case could be avoided if the parameter of interest were restricted to a finite number of values in m intervals on the real line, and if m were then allowed to tend to larger values:

> [§]25... Suppose, for instance, that a point lies on a line of length $m.l.$, we may write the alternative 'the interval of length l on which the point lies is the xth interval of that length as we move along the line from left to right' $\equiv f(x)$; and the principle of indifference can then be safely applied to the m alternatives f(1),

> f(2),...,f(m), the number m increasing as the length l of the intervals is diminished. There is no reason why l should not be of any definite length however small. (1921 [1973], pp. 67-68)

It is unclear how this will help. We must first determine which value of m we take to be indivisible: surely any such choice will be arbitrary; indeed, it would seem to be just wrong. Second, it seems that if m does tend to infinity, we once again are dealing with a continuously valued parameter, and so the paradoxes will reemerge. Also, it is difficult to interpret Keynes on this matter, because he seems to contradict himself. He first says that the Principle of Indifference is not applicable in the continuous case (§24), and then immediately after this (§25) says that it is.

3. Howson and Urbach on Keynes's solution

Howson and Urbach (1993) interpret Keynes as saying that we should apply the Principle of Indifference to some number, m of equal intervals of values of a certain variable, and then let m tend to infinity. That is, we apply the Principle of Indifference for each value of m as m goes to infinity. This seems to amount to presupposing a uniform probability distribution over the variable (in the example above, we would only allow $\frac{d\phi}{b-a}$ as the density distribution for the parameter θ). The density distribution would not be invariant under a large class of transformations, but since we presuppose a uniform probability distribution, the transformations would be, *ex hypothesi*, ruled out. This, Howson and Urbach interpret Keynes as saying, is in keeping with practice in mathematical physics, "where often the form of a density distribution depends on the passage to the limit." (Howson and Urbach 1993, p. 61. They are referring to Keynes's remark quoted on page 21.)

As Howson and Urbach note, this is hardly a solution: it requires an "arbitrary choice of how in any given case to proceed to the limit (that is, which particular variable to take as determining the uniform distribution)." (Howson and Urbach, *ibid*) Keynes might reply to the charge of arbitrariness that there are some "natural" predicates or variables. Although Keynes does not seem to have suggested this, it would be in keeping with his views on intuition. However, there seems little prospect for any such proposal: A definition of "natural" in this context is highly problematic. First, it is not clear why "naturalness" should confer any special status on a predicate as regards the Principle of Indifference. And, it seems that any definition of naturalness is arbitrary in that it would be dependent on our current scientific knowledge. Further, there seems no reason that for a "natural" variable S, there might be some function f(S) which is also natural, but which gives a different distribution after the Principle of Indifference is applied. This does seem to be the case with velocity and velocity squared (which is proportional to kinetic energy). Finally, it is not clear restricting the application of the Principle to certain predicates could solve Bertrand's paradox. Even if we were to apply the Principle as Keynes recommends, it is not clear that we would not still derive different probabilities for the wine/water and water/wine ratios (of course, unless we do not apply the Principle to any quantity which can be inverted, which seems to rule out *all* parameters). I conclude that if Howson and Urbach are correct in their interpretation of Keynes, his version of the Principle of Indifference is fatally flawed.

However, Howson and Urbach's interpretation seems to contradict the last line of the above, admittedly ambiguous, quote ("There is no reason why l should not be of any definite length however small"). If m tends to infinity, then l would not be "an indefinite length", but infinitesimal. Hence it might be useful to look for another interpretation of Keynes's solution: this is given by Gillies.

4. Gillies on Keynes's solution

According to Gillies (1973, pp. 12-13, 1988, p. 187), Keynes disallows the Principle of Indifference in the continuous case by disallowing real-valued parameters. Instead, real-valued parameters should be considered as parameters which can take a large but not denumerable number of values. In other words, Gillies interprets Keynes's solution to be that at some stage we must stop dividing a continuous interval into smaller sub-intervals. This does seem to be in keeping with Keynes's remark that "There is no reason why l should not be of any definite length however small." However, as Gillies notes, this contradicts Keynes's restriction that the Principle be only applied in cases where the alternatives are not further divisible. Further, it does not solve the problem. As Gillies shows, we can reconstruct Bertrand's paradox even when we have a large, finite number of subdivisions. Consider a partition of the interval [1/6,3] into intervals $[a,b]$ of, say, length 1/6. According to the Principle of Indifference, the probability that the ratio falls in any one of the intervals is 1/16. In this case p(wine/water$\leq$2) = 10/16 and p(water/wine$\geq$1/2) = 15/16, which is a contradiction. If we further decrease the length of the interval to, say, 1/12, we still get the contradiction, and for any subintervals of length $(1/6)2^{-n}$, or indeed for the general case of intervals of length $(1/x)y^{-n}$:

> Suppose in the wine-water example we divide the interval (1/3,3) into n equal subintervals $I_1,\ldots,I_n$ and consider the event E that there is less than twice as much wine as water. By taking the length of I_i sufficiently small and representing E first as a combination of events of the form wine/water $\in I_j$ and

> then of events of the form water/wine $\in I_k$, we obtain by suitably modifying the previous argument two different probabilities for E. (Gillies 1973, p. 13)

Finally, as Gillies points out, this solution makes the Principle inapplicable in most scientific procedures, since much of science uses continuous parameters:

> ...the continuous parameter case is the one most commonly met with in scientific practice. Whatever the area of investigation, it is normal to introduce explanatory hypotheses which depend on a finite number of continuous parameters... Thus Keynes solves the paradoxes of the Principle of Indifference only by rendering the logical theory of probability quite unsuitable for use with most scientific hypotheses. (Gillies 1988, p. 187)

I conclude that if Gillies's interpretation is correct, then once again, Keynes's version of the Principle of Indifference leads to insurmountable difficulties. But, in the next section, I shall also express doubts about Gillies's interpretation.

5. An alternative interpretation of Keynes's solution

There is reason to doubt both Howson and Urbach's and Gillies's interpretation of Keynes. It would be strange indeed if Keynes had not noticed that he immediately contradicted himself with respect to the criterion that alternative hypotheses must not be further divisible. In fact, Keynes seems to have rejected what Gillies interprets as his solution, although for a different reason: Keynes held that applying the Principle of Indifference to some set of very small sub-intervals was just as legitimate (or not) as applying the Principle to the original intervals. Keynes discussed a proposal to apply the Principle of Indifference only to the smallest discernable divisions (that is, the limits our of technological capabilities would determine how we apply the Principle): "A. Nitsche... in criticising von Kries, argues that the alternatives to which the principle must be applied are the smallest *physically* distinguishable intervals, and that the probability of the specific volume's lying within a certain range of values turns on the *number* of such distinguishable intervals in the range." (Keynes, 1921 [1973], p. 49, n.1) However, since Keynes interpreted probability as objective, he wished to provide only one single probability for any given statement. Relying on our technology to provide a basis for probability assignments violates objectivity, since it is dependent on contingent factors:

> This procedure might conceivably provide the correct method of computation, but it does not therefore restore the credit of the principle of indifference. For it is argued, not that the results of applying the principle are always wrong, but that it does not lead unambiguously to the correct procedure... even if we... reckon intervals as equal which contain an equal number of 'physically distinguishable' parts, is it certain that this does not simply provide us with a new system of measurement, which has the same conventional basis as the methods of specific volume and specific density, and is no more the one correct

> measure than these are? (ibid, p. 49, n.1)

(Keynes is alluding to the paradoxes like the wine/water one which arises from a transformation of the variables.) Elsewhere, Keynes seems to have conceded that the Principle is inapplicable in most continuous cases: "M. Bertrand is so much impressed by the contradictions of geometrical probability that he wishes to exclude all examples in which the number of alternatives is *infinite*. It will be argued in the sequel that something resembling this is true." (ibid, p. 53)

I propose to interpret Keynes differently than Gillies and Howson and Urbach. Consider the paradoxes in the continuous case: Keynes regarded the paradoxes of geometric probability as significantly different from the paradoxes exemplified by the wine/water example. First he deals with these two paradoxes in two different sections—the wine/water in §23 (pp. 66-67), and the geometric in §25 (pp. 67-68). Second, he seems quite explicit that his solution to the geometric paradox (which he expounds in §25) does not apply to paradoxes of the wine/water type (in §23):

> The application of this result [Keynes's reformulated Principle of Indifference]... deals with the examples of ranges of specific volume and specific density, because there is no range which does not contain within itself two similar ranges. As there are in this case no definite units by which we can define *equal* ranges, the device, which will be referred to in §25 for dealing with geometrical probabilities, is not available. (1921 [1973], p. 66-67)

This throws light on Keynes's discussion immediately following (§24 quoted above). He is clearly saying that the Principle of Indifference does not apply to the wine/water case, and so he does not claim to have shown how to overcome the paradoxical results of the application of the Principle of Indifference in this case.

The solution proposed in §25 (quoted above "Suppose, for instance...") is only to be applied to geometric probability, and not to the determination of the probability of a point taking a particular position on a continuum: "If we are dealing with a strictly linear chord, the Principle of Indifference would yield us no result, as we could not enunciate the alternatives in the required form..." (1921 [1973], p. 68). However, Keynes's solution is obscurely phrased. He seems to regard each of the three alternatives (figures 2a,b,c) as representing different shapes, "distinct elementary areas", and so as representing three different problems:

> If we deal with the problems of geometrical probability in this way, we shall avoid the contradictory conclusions, which arise from confusing together *distinct* elementary areas. In the problem, for instance, of the chord drawn at random in a circle... the chord is regarded, not as a one-dimensional line, but as the limit of an area, the shape of which is different in each of the variant solutions. In the

> first solution it is the limit of a triangle, the length of the base of which tends to zero; in the second solution it is the limit of a quadrilateral, two of the sides of which are parallel and at a distance apart which tends to zero; and in the third solution the area is defined by the limiting position of a central section of undefined shape. (1921 [1973], p. 68)

Exactly how Keynes meant to overcome Bertrand's paradox is not clear. He says the chord is to be represented as the limit of an area as it tends to zero: But in what Keynes calls the "first solution" (figure 2a), in which "the length of the base... tends to zero", the area of the triangle also tends to zero. The area does, however, tend to a line, which has zero area. Whatever Keynes envisaged as a solution, it is clear that he meant to regard the three separate cases of the geometric paradox as representing three different problems. But this is clearly wrong: since we can determine from the different shapes the lengths of the associated chords, we still get three different answers concerning the length from the Principle of Indifference. It seems, once again, that Keynes's proposal amounts simply to ruling out certain transformations: In particular, he is ruling out non-linear transformations of coordinate systems (Neyman 1952, pp. 15-17, represents Bertrand's paradox in terms of alternative coordinate systems, Howson 1976, pp. 287-288, gives the transformations used). Keynes's reasons for ruling out these transformations is that this is sometimes done in physics, where the determination of the value of a variable represented as a limit depends on how the limit is reached:

> The substance of this explanation can be put in a slightly different way by saying that it is not a matter of indifference in these cases in what manner we proceed to the limit. We must assign the probabilities *before* proceeding to the limit, which we can do unambiguously. But if the problem in hand does not stop at small finite lengths, areas, or volumes, and we have to proceed to the limit, then the final result depends upon the shape in which the body approaches the limit. Mathematicians will recognise an analogy between this case and the determination of potential at points *within* a conductor. Its value depends upon the shape of the area which in the limit represents the point. (1921 [1973], pp. 68-69)

This response does not help Keynes, as we noted above, because he must rule out transformations of random variables solely to preserve the consistency of the Principle of Indifference, and this has disastrous consequences for the practical application of the Principle. In conclusion, I disagree to some extent with the interpretation of Howson and Urbach, and Gillies, because they do not separate Keynes's solution to the paradox into the continuous and the geometric. Nonetheless, their criticisms still apply, and therefore it seems that Keynes has not solved the problems of the Principle in the continuous case.

Finally, it is not clear that Keynes solved the problems of the Principle of Indifference even in the discrete case. Colour can be represented as a continuous term (as a point in the continuum of electromagnetic radiation), as can the faces of a die (as surfaces of a cube in R^3), and so propositions which include these terms will raise the same difficulties. (This also shows that our representation of the variables to which the Principle of Indifference is to be applied is language dependent, and to this degree, not objective.) I conclude that Keynes's modified Principle of Indifference is not a valid constraint on probability distributions and that, hence, his programme for a probabilistic methodology is a failure. I will now examine the implications of this failure for Keynes's account of justification.

C. The Influence of Moore's Ethical Epistemology on Keynes's Epistemology

As we have seen, Keynes was unsuccessful in his attempt to formulate a consistent and adequate version of the Principle of Indifference. I wish to examine the implications of this failure for Keynes's intuitionistic approach. I will argue that Keynes's foundations of probability are very similar to G.E. Moore's foundations of ethics, and that they share the same, insurmountable, difficulties. Thus in section 1, I explicate G.E. Moore's account; in section 2, I argue that this account faces insurmountable difficulties; in section 3, I show that Keynes's and Moore's accounts are in crucial respects the same, but applied to different subject matters (ethics and probability); and in section 4, I conclude that Keynes's attempt to found probability on the basis of intuitions is irredeemably flawed, as must be all such attempts.

Keynes was strongly influenced by Moore's theory of ethics. In the preface to *A Treatise on Probability* Keynes acknowledges this influence: "It may be perceived that I have been much influenced by W.E. Johnson, G.E. Moore, and Bertrand Russell, that is to say by Cambridge..." (1921 [1973], preface) In fact, in his essay "My Early Beliefs" Keynes attributed his interest in probability in part to Moore:

> He [Moore] also has a section on the justification of general rules of conduct. The large part played by considerations of probability in his theory of right conduct was, indeed, an important contributory cause to my spending all the leisure of many years on the study of that subject: I was writing under the joint influence of Moore's *Principia Ethica* and Russell's *Principia Mathematica*. (Keynes 1938 [1972], p. 445)

Keynes goes on to say that he accepted neither Moore's view of probability nor his account of the relation of probability to rules of conduct. In particular, Keynes regarded Moore's frequentist interpretation of probability as untenable. His disagreement with Moore's interpretation of probability seems to have been related, or perhaps motivated by his rejection of Moore's conclusions about the ethical implications of an intuitionistic approach. In *Principia Ethica*, Moore argued that a society's moral rules served as a rational guide to right action because these rules, if followed, would with high probability lead to right actions. Moore concluded that an individual was never justified in exempting himself from such a moral rule:

> ...Can the individual ever be justified in assuming that his is one of these exceptional cases [in which "an established rule" may be violated]? ...it seems that this question may be definitely answered in the negative. For, if it is certain that in a large majority of cases the observance of a certain rule is useful, it follows that there is a large probability that it would be wrong to break the rule in any particular case; and the uncertainty of our knowledge both of effects and of their value, in particular cases, is so great, that it seems doubtful whether the individual's judgment that the effects will probably be good in his case can ever be set against the general probability that that kind of action is wrong. (Moore 1903 [1968], p. 162)

Keynes rejected Moore's frequentist account of probability, and hence Moore's argument for always following an established rule. Since Keynes held that the probability of a rule's leading to a good result could only be determined by intuition, which was a matter of what an individual perceived, the individual was not necessarily under any obligation to follow established moral rules:

> ...we set on one side, not only that part of Moore's fifth chapter on 'Ethics in relation to Conduct' which dealt with the obligation so to act as to produce by causal connection the most probable maximum of eventual good through the whole procession of future ages (a discussion which was indeed riddled with fallacies), but also the part which discussed the duty of the individual to obey general rules. We claimed the right to judge every individual case on its merits, and the wisdom, experience and self-control to do so successfully... We repudiated entirely customary morals, conventions and traditional wisdom. (Keynes "My Early Beliefs" 1938 [1972], p. 446)

There has been much debate (see, e.g., Skidelsky 1983) as to the accuracy of Keynes's report of Moore's influence on him, and of the accuracy of Keynes's report of the ethical views of his and the Bloomsbury group's ethical views, in "My Early Beliefs." An account, drawn from Keynes's unpublished papers, of Moore's considerable influence may be found in Moggridge 1992 and Skidelsky 1983. However, it seems clear from Keynes's many statements of Moore's influence on him, in both his published and unpublished writings, that although Keynes disagreed with much in Moore, and developed his theory of probability in opposition to Moore's, Keynes adopted Moore's intuitionistic account of justification. And the fact that Keynes discussed his then unpublished *Treatise on Probability* with Moore (Moore even read the proofs) suggests Moore's strong influence on this work: but, unfortunately, we have no details of their discussions, only the dates of appointments (Bateman 1988).

1. Moore's account of justification

I will establish that Keynes's attempt to justify the foundations of probability followed the same lines as Moore's attempt to justify the theory of ethics. But before I can do so, it is necessary to examine Moore's views on how ethical statements are justified, and the difficulties these encounter (this will occupy the next two sections).

Moore's theory is usually labelled "intuitionism", since he held that we know that something is good because we intuit, or perceive in some way, the property "good," which it supposedly possesses. The property of goodness, according to Moore, is not definable in terms of other properties because it is itself ultimate or 'simple.' He also held that 'good' was a non-natural property. Moore concluded that although we may come to know that something is 'good', we cannot do so by defining good in terms of physical properties and then observing whether that thing has those physical properties. I will examine each of these claims in turn.

Moore distinguished complex from simple properties. Complex properties may be defined in terms of simpler properties: Moore used the property 'horse' as an example. You can define 'horse' as a kind of mammal, which can then be defined in further simpler properties—but if someone is unfamiliar with the simplest properties of which these are composed, they will not be able to understand 'horse':

> You can give a definition of a horse, because a horse has many different properties and qualities, all of which you can enumerate. But when you have enumerated them all, when you have reduced a horse to his simplest terms, then you can no longer define those terms. They are simply something which you think of or perceive, and to any one who cannot think of or perceive them, you can never, by any definition, make their nature known. (Moore 1903 [1968], p. 7)

Moore thus seems to divide properties into those that can be defined in terms of others, and those which may only be defined ostensively (this seems to be the meaning of the last sentence of the above quote: "They are simply something which you think of or perceive, and to any one who cannot think of or perceive them, you can never, by any definition, make their nature known"). According to Moore, "good" is a simple property, which cannot be defined.

However, one might argue against more that simple properties could be correlated with other, perhaps complex, physical properties, and thus defined. This is, according to Moore, not possible, since, he claimed, there is a distinction between so called 'natural' and 'non-natural' simple properties. This is perhaps best illustrated by Moore's example of the property "yellow." If you try to explain to someone who is unfamiliar with yellow what it is, then all you can do is give an ostensive definition, that is, point to it. If they are still unable to understand what you mean by 'yellow', then you can do nothing else to communicate the meaning of the term to them. Moore considered the possibility that "yellow" might be a complex property: one could try to define a property like yellow by means of physical properties such as frequencies of a certain type of radiation which causes human visual systems to register 'yellow.' Moore denied this possibility: he claimed that there was a readily perceivable distinction between "yellow" as defined in terms of physical properties (this is presumably the natural property "yellow"), and "yellow" as perceived by the mind's eye:

> Consider yellow, for example. We may try to define it, by describing its physical equivalent; we may state what kind of light-vibrations must stimulate the normal eye, in order that we may perceive it. But a moment's reflection is sufficient to shew that those light-vibrations are not themselves what we mean by yellow. *They* are not what we perceive. Indeed we should never have been able to discover their existence, unless we had first been struck by the patent difference of quality between the different colours. The most we can be entitled to say of those vibrations is that they are what corresponds in space to the yellow which we actually perceive. (ibid, p. 10)

Moore famously also classified "good" as a non-natural, simple property, and so, like "yellow", Moore held that "good" could not be defined: "If I am asked 'What is good?' my answer is that good is good, and that is the end of the matter. Or if I am asked 'How is good to be defined?' my answer is that it cannot be defined, and that is all I have to say about it." (1903 [1968], p. 6). But, of course, he did have more to say about the matter. Good, he held, is undefinable for the same reasons that 'yellow' is—it is a simple, non-natural property:

> My point is that 'good' is a simple notion, just as 'yellow' is a simple notion; that, just as you cannot, by any manner of means, explain to anyone who does not already know it, what yellow is, so you cannot explain what good is. Definitions of the kind that I was asking for, definitions which describe the real nature of the object or notion denoted by a word, and which do not merely tell us what the word is used to mean, are only possible when the object or notion in question is something complex. (ibid, p. 7)
>
> 'Good,' then, if we mean by it that quality which we assert to belong to a thing, when we say that the thing is good, is incapable of any definition, in the most important sense of that word. The most important sense of 'definition' is that in which a definition states what are the parts which invariably compose a certain whole; and in this sense 'good' has no definition because it is simple and has no parts. It is one of those innumerable objects of thought which are themselves incapable of definition, because they are the ultimate terms by reference to which whatever *is* capable of definition must be defined... There are many instances of such qualities. (ibid, pp. 9-10)

Moore held that it was a mistake to define properties like good and yellow by reference to their "physical equivalents"; he famously termed this the "the naturalistic fallacy":

> It may be true that all things which are good are *also* something else, just as it is true that all things which are yellow produce a certain kind of vibration in the light. And it is a fact, that Ethics aims at discovering what are those other properties belonging to all things which are good. But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were simply not 'other', but absolutely and entirely the same with goodness. This view I propose to call the 'naturalistic fallacy'... (ibid, p. 10)

Moore's distinction between natural and non-natural properties is notoriously difficult to understand. Briefly stated, it is the distinction between those objects studied by the sciences, which, according to Moore, are things which have existed, do exist, or will exist "in time", and others:

> By 'nature,' then, I do mean and have meant that which is the subject-matter of the natural sciences and also of psychology. It may be said to include all that has existed, does exist, or will exist in time. If we consider whether any object is of such a nature that it may be said to exist now, to have existed, or to be about to exist, then we may know that that object is a natural object, and that nothing, of which this is not true, is a natural object. (ibid, p. 40)

This definition leaves unclear which properties are included in the subject matter of the natural sciences: Moore rejected a naturalistic definition of 'yellow'. Perhaps Moore held that the subjective experience of 'yellow' was non-natural, while the "physical equivalent" of yellow was natural. This seems likely, since Moore held that non-natural properties could not be conceived of as existing "in time" independently of objects:

> ...I do not deny that good is a property of certain natural objects: certain of them, I think, *are* good; and yet I have said that 'good' itself is not a natural property. Well, my test for these too also concerns their existence in time. Can we imagine 'good' as existing *by itself* in time, and not merely as a property of some natural object? For myself, I cannot so imagine it, whereas with the greater number of properties of objects—those which I call the natural properties—their existence does seem to me to be independent of the existence of those objects. (ibid, p. 41)

Moore's distinction between natural and non-natural properties is not very satisfactory. Perhaps he meant something as follows: We can imagine in the mind's eye a yellow patch not attached to any object, but it seems that we cannot imagine a yellow patch in the physical world, unless it is the colour of some object. Likewise, with good: good would seem always to accompany states of affairs, and never occur on its own.

Finally, Moore held that there were four classes of "goods" and "evils": (1) Unmixed goods; (2) Unmixed evils; (3) mixed goods; and (4) mixed evils. (3) and (4) are things which, while predominately good or evil, contain some elements of both. (1) is, according to Moore "the love of beautiful things or of good persons", (2) is "either (a) ...the love of what is evil or ugly, or (b) ...the hatred of what is good or beautiful, or (c) ...the consciousness of pain." (ibid, pp. 224-225)

In summary, Moore held that there were certain properties which by their very nature were not amenable to scientific investigation: they can be merely pointed to, and never described in terms of other objects. His test for which properties are of this sort is, at best, difficult to understand. The next section is devoted to criticising his account.

2. Criticism of Moore's intuitionism

It may seem that "yellow" and "good" are transparent properties. However, under an intuitionistic account this is not so: they are, in fact, inherently mysterious (quite unlike what one would expect). As we shall see, this feature leads to serious difficulties for the intuitionistic account. I will survey some of these difficulties in this section. I will then argue Keynes's intuitionistic account of probability (developed in section 3) encounters similar difficulties, although in a more extreme manner (section 4).

Moore offers two reasons why we cannot define good in empirical terms. First, that the non-natural character of 'good' (determined by its independent existence or not "in time") precludes the possibility of stating any detectable (external) conditions under which it occurs. Second, the simple character of 'good' means that it cannot be studied by breaking it down into, and explaining it in terms of, other, perhaps more familiar terms, but instead, if it is to be discussed at all, it must merely pointed to. The first claim has been strongly criticised (see, for example Frankena 1963, Warnock 1967, Feldman 1978 and MacIntyre 1984), mainly on the grounds that intuitionism leaves unexplained how the supposedly intuited non-natural properties are related to physical properties, which, unless we wish to assume the existence of mental substances, is necessary; this seems to imply that a particular non-natural property could occur in one, but not in another, physically identical, circumstance. Moreover, intuitionism makes rational debate concerning the presence of an intuited property impossible. Also, the presence of a non-natural property can provide no reason for changing our behaviour. Finally, there just do not seem to be such properties as Moore describes. I will take these criticisms in turn.

The claim that 'good' is non-natural implies that it is not related to the physical features of the world—if it were, we would be able to state the physical conditions under which it appears. Moore explicitly rejected any link between a non-natural property and any "physical equivalent." The only way in which a non-natural property can be detected is through our putative powers of intuition, not through scientific investigation. But if this were the case, then we could give no reasons why some situations are good, and others not. G.J. Warnock makes this objection:

> The picture presented is that of a realm of moral qualities, *sui generis* and indefinable, floating, as it were, quite free from anything else whatever, but cropping up here and there, quite contingently and for no reason, in bare conjunction with more ordinary features of the everyday world. (Warnock 1967, p. 14)

Since non-natural properties are inexplicable, it is difficult to see how anything is 'good.' Presumably, Moore would have to claim that humans possess a faculty which allows us to 'see' good. But this faculty must be unconnected with our brain processes, otherwise it would be possible, in principle, to analyse good in terms of the physical processes which lead to its detection. Since, under this account, good is independent of any physical property, it permits the possibility that given two physically identical situations, with 'good' occurring in one and not the other. (This is sometimes referred to as the problem of supervenience.) Such a result highlights the peculiarity of the intuited property, and the implausibility of this view.

Second, if the property 'good' is independent of natural properties, its presence or absence does not seem to provide a motivation for moral behaviour, and seems to imply that we are under no obligation to act according to moral standards: the presence of good would make no difference to the world. Thus, even if Moore were correct that the property 'good' is non-natural, this has no consequences for our actions. If good is connected with natural properties (say, pleasure), then we may argue that presence provides motivation for our acting so as to bring about its presence.

Third, this type of non-naturalism makes debate about whether a situation possesses a non-natural property or not impossible. However, it seems that debate about 'good' is possible. Suppose that two persons, X and Y, disagree as to whether something is good. According to Moore's intuitionism, one of them must be wrong, that is, one of them must be misperceiving, or morally blind. Suppose person Y wishes to show X that he is wrong, or wishes to check whether he himself is wrong. Y will not be able to appeal to any criteria to settle this dispute, because the only criterion is whether or not 'good' has been perceived. 'Good' is, according to intuitionism, not necessarily accompanied by any other properties which could help us to establish its presence. Therefore, the only option available is to assert its presence if you feel you have intuited it. No debate is possible between persons who disagree. Perhaps it is an inevitable feature of moral questions that agreement cannot be forced (I will argue later that in the case of probability we have good reason for not thinking that this is a feature of debate about the nature of probability).

For Moore, this could be the case with 'natural' properties. Moore held that the property 'yellow' could not be identified with light of some particular frequency or frequencies. But if we wish to check if something is 'yellow', then one way of doing this is surely to check the frequency of the reflected light, and that our perceptual apparatus is working correctly. If all is in order, then it seems we would be able to give a confident answer as to whether an object is yellow or not. Thus, there seems to be an objective, or intersubjective criterion for settling disputes about what someone is perceiving or not. But perhaps Moore would deny this, as would philosophers of mind who hold that spectrum inversion is possible (that is, there could be persons who perceive colours differently, say red where everyone else sees green, but whose behaviour and responses in relation to colours are indistinguishable from those who perceive colours differently).

This leads to the objection that the inability to force agreement seems to indicate that there are no such properties. Many, myself included, feel that even after introspection, the kinds of properties that Moore describes do not seem to exist. (Certainly Moore's test for the naturalness of a property does not seem to be intuitively obvious, nor does his account of what is good.) The arguments discussed so far do not establish intuitionism to be false, but merely implausible. But, if an alternative account were available, intuitionism would be unnecessary. So the arguments discussed so far do give impetus to the search for an alternative. This conclusion is reached by Frankena: "On the whole... intuitionism strikes me as unplausible even if it has not been disproved... The main point... is that the belief in self-evident ethical axioms and value-judgments, and all that goes with it, is so difficult to defend that it seems best to look for some other answer to the problem of justification." (Frankena 1963, p. 88) MacIntyre makes a much stronger claim: the failure of the intuitionist to demonstrate the existence of non-natural properties is the same reason we no longer believe in witches and unicorns. In both cases, persistent attempts to demonstrate their existence have failed. (Actually, MacIntyre makes this claim about the existence of human rights, but the point applies equally well to intuitionism):

> ...the truth is plain: there are no such [human] rights, and belief in them is one with belief in witches and unicorns.
>
> The best reason for asserting so bluntly that there are no such rights is indeed of precisely the same type as the best reason which we possess for asserting that there are no witches and the best reason which we possess for asserting that there are no unicorns: every attempt to give good reasons for believing that there *are* such rights has failed. The eighteenth-century philosophical defenders of natural rights sometimes suggest that the assertions which state that men possess them are self-evident truths. Twentieth-century moral philosophers have sometimes appealed to their and our intuitions; but one of the things that we ought to have learned from the history of moral philosophy is that the intro-

duction of the word 'intuition' by a moral philosopher is always a signal that something has gone badly wrong. (MacIntyre 1984, p. 69)

Further, it seems that an alternative to intuitionism is available, namely, that we just identify those putative non-natural properties their physical correlates. Moore acknowledges that just as the perception of the colour yellow is always accompanied by the reception of light of a certain frequency (and presumably certain types of physiological activity), so might good situations always be accompanied by certain physical events. But if there were a universal correlation between a simple property and certain physical events, then nothing would seem to be lost by identifying the property and the physical events. No doubt Moore would disagree on the grounds that even if such a scientific explanation were available, it would fail to capture what 'good' meant. But the qualities omitted by the physical definition must be of a peculiar sort, as noted above. It just may be the case that phenomenological and ethical properties are very peculiar, and this no doubt is why many would agree that water is the same as H_2O, but disagree that the mind is the same as the brain. (However, I shall argue that there is no reason to think that "probable" is such a peculiar property.)

The second claim, that 'good' is simple and hence indefinable, has some plausibility. Consider Moore's example of 'yellow.' We may perhaps discover that yellow is associated with a certain frequency (5.2nm) in the electromagnetic spectrum. But, as Moore pointed out, a blind person can know the frequency and indeed detect the frequency associated with the colour yellow, and still not know what it is like to see yellow. This argument, however, rests on the notion of qualia, that we have irreducibly subjective experiences, and further, that these irreducibly subjective experiences are in fact not explainable in terms of other properties (even other subjective experiences). This is, of course, an issue of much contention in contemporary philosophy of mind, and I cannot claim to be able to settle it. But I will argue later that no matter what plausibility this view may have in ethics and the philosophy of mind, it has very little in the philosophy of science, and particularly as an account of the foundations of probability.

However, an example of why we should be wary of intuitions of the simplicity of *certain* properties is provided by psychological studies of colour perception. Richard Gregory took it as unfortunate for Moore that he chose the example of 'yellow' as a simple property since it is "physiologically, always a mixture of red-green neural signals." (Gregory 1987, p. 496) Gregory draws the conclusion that:

> This is a warning that it is not possible to infer the simplicity of physiological mechanisms from the simplicity or apparent complexity of experience. It shows indeed the fallibility of introspection for understanding functions of the nervous

system, if not of the mind. (Gregory 1987, p. 818)

Moore might respond that Gregory has missed the point—yellow may be physiologically complex, but it is, Moore would claim, phenomenologically simple. On the other hand, experiments described by Dennett seem to show that the phenomenology of seemingly simple colours like yellow is anything but simple: For example, if a subject is shown a yellow circle on a blue patch, and then the brightness of the colours is equalised, the boundary between the two colours disappears. (Dennett 1991, p. 69) This type of consideration undermines the power of intuitions to determine whether or not a property is 'simple': our intuitions are shown to be wrong by empirical investigation, and hence inapplicable. I will argue in the case of scientific method that the domain of judgments which could be settled by appeals to intuition is negligible. I conclude that the objection that non-naturalistic properties are inherently mysterious, so mysterious that they do not seem to exist, has considerable force.

I will argue that in the case of the philosophy of science, the problems discussed are much more difficult, and in fact fatal to intuitionism, and so more strongly motivate the search for an alternative.

3. Keynes's theory of justification

I will argue in this section that Keynes's view of the foundations of probability is the same as Moore's account of the foundations of ethics. Keynes's debt to Moore has only been noticed in recent years, despite Keynes's explicit acknowledgment of Moore's influence. It seems that, for the most part, Keynes's account of probability is the same as Moore's account of "good." A.M. Carabelli makes this point: "Keynes's probability shared all the attributes of Moore's concept of goodness: it was a simple notion, unanalysable, indefinable, non-natural, directly perceived or intuited and objective." (1988, p. 31) Moggridge (1992, pp. 148-152) agrees, as does O'Donnell (1991, p. 9). As I argued in section A, Keynes held that probability was a simple, indefinable and directly intuited property. However, Keynes never states explicitly that probability is non-natural, although it seems plausible, as Carabelli says, that he did view it as such.

I disagree with Carabelli's interpretation on one point: she claims that Keynes disagreed with Moore as to the absolute validity of intuitive judgments. Carabelli argues that, in contrast to Moore's account of the intuition of "good," Keynes held that intuitions of probability relations were objective but not necessarily "universal and self-evident":

> ...Keynes, though rejecting the subjective and individualist vision of the relation of probability, did not accept its opposite representation as universal and self-evident. In other words, the attribute of 'objectivity' did not automatically carry

> with it, as in Moore, the characters of universality and self-evidence. (Carabelli 1988, p. 38)

Carabelli's claim that, according to Keynes, intuitive judgments are not necessarily "universal" may be interpreted in two ways. First, probability relations are objective in the sense that they have an independent existence, but may only be intuited when certain favourable conditions occur; or second, that persons who disagree on the probability of a hypothesis may both be correct, because probabilities are not "universal." I will argue that the first position is Keynes's, and that Keynes did not significantly differ from Moore on this point. And, I argue that the second position is incoherent, and was not adopted by Keynes.

Keynes explicitly acknowledged that our intuitions of logical relations are not always perfect, and that some are better able to intuit logical relations than others: "Some men—indeed it is obviously the case—may have a greater power of logical intuition than others." This uneven distribution of intuitive powers may lead to disagreements about what is self-evident: "What is self-evident to me and what I really know, may be only a probable belief to you, or may form no part of your rational beliefs at all. And this may be true not only of such things as *my* existence, but of some logical axioms also." Keynes also agreed that intuition may not capture all logical relations: "We can no more assume that all true secondary propositions are or ought to be known than that all true primary propositions are known. The perceptions of some relations of probability may be outside the powers of some or all of us." Keynes concluded that there does seem to be an element of subjectivity in our intuitions of probability relations: "What we know and what probability we can attribute to our rational beliefs is, therefore, subjective in the sense of being relative to the individual." (Keynes 1921 [1973], pp. 18-19)

However, Keynes did not think that the limits on our powers of intuition which he described were universal, and he held that at least some people were able to intuit correctly some logical relations and premises, and that this was sufficient to construct a logical probability relation. He continues after the above quotes:

> *But given the body of premises which our subjective powers and circumstances supply to us*, and given the kinds of logical relations, upon which arguments can be based and which we have the capacity to perceive, the conclusions, which it is rational for us to draw, stand to these premises in an objective and wholly logical relation. Our logic is concerned with drawing conclusions by a series of steps of certain specified kinds from a *limited* body of premises. (Keynes 1921 [1973], p. 19, italics mine)

Thus Keynes believed that we could make objectively valid probabilistic arguments. This interpretation is bolstered by Keynes's statements made later in the *Treatise*, where he acknowledges once again that powers of intuition may fail in some cases:

> To say, then, that a probability is unknown ought to mean that it is unknown to us through our lack of skill in arguing from given evidence. The evidence justifies a certain degree of knowledge, but the weakness of our reasoning power prevents our knowing what this degree is. At the best, in such cases, we only know *vaguely* with what degree of probability the premises invest the conclusion. The [sic (presumably "that")] probabilities can be unknown in this sense or known with less distinctness than the argument justifies, is clearly the case. We can through stupidity fail to make any estimate of a probability at all, just as we may through the same cause estimate a probability wrongly. (ibid, pp. 34-35)

But Keynes then refers to his discussion quoted above, and makes the same point: that although we can intuit only some logical relations, this does not mean that there is no objective probability relation:

> But this admission must not be allowed to carry us too far. Probability is, *vide* chapter 2 [§11 quoted above], relative to *human* reason. The degree of probability, which it is rational for *us* to entertain, does not presume perfect logical insight, and is relative in part to the secondary propositions which we in fact know; and it is not dependent upon whether more perfect logical insight is or is not conceivable. It is the degree of probability to which those logical processes lead, of which our minds are capable; or, in the language of chapter 2, which those secondary propositions justify, which we in fact know. (ibid, p. 35)

However, Keynes seems to argue fallaciously for his view that we can intuit sufficiently many probability relations to construct an inductive logic as follows: If we do not hold that we can intuit probability relations correctly, then we cannot construct an inductive logic. Therefore, Keynes seems to conclude, we must be able to intuit probability relations:

> If we do not take this view of probability [presumably the view of probability as objective and intuitable within the constraints of human reasoning], if we do not limit it in this way and make it, to this extent, relative to human powers, we are altogether adrift in the unknown; for we cannot ever know what degree of probability would be justified by the perception of logical relations which we are, and must always be, incapable of comprehending. (ibid, p. 35)

Nonetheless, I think that these passages (and the passage quoted on p. 4) make clear that Keynes held that probability relations were objective in the sense that between any two propositions there was ~~only~~ one and only one probability relation.

I will now argue that if probability relations are not taken to be objective in this sense (the second possible interpretation of Carabelli's position) then Keynes's position would be incoherent. Ramsey interpreted Keynes as holding that probability relations were not objective in the sense that only one probability relation held between any set of propositions. Ramsey argued that to hold that logical intuition was not objective was to give up intuitionistic foundations for probability:

> Another argument against Mr. Keynes' theory can, I think, be drawn from his inability to adhere to it consistently even in discussing first principles. There is a passage in his chapter on the measurement of probabilities which reads as follows: [Ramsey then quotes the two paragraphs from Keynes given immediately above]...
>
> This passage seems to me quite unreconcilable with the view which Mr. Keynes adopts everywhere except in this and another similar passage [probably the one quoted above from chapter 2]. For he generally holds that the degree of belief which we are justified in placing in the conclusion of an argument is determined by what relation of probability unites that conclusion to our premises. There is only one such relation and consequently only one relevant true secondary proposition, which, of course, we may or may not know, but which is necessarily independent of the human mind. If we do not know it, we do not know it and cannot tell how far we ought to believe the conclusion. But often, he supposes, we do know it; probability relations are not ones which we are incapable of comprehending. But on this view of the matter the passage quoted above has no meaning: the relations which justify probable beliefs are probability relations, and it is nonsense to speak of them being justified by *logical relations which we are, and must always be, incapable of comprehending*. (Ramsey 1978, pp. 65-66. I have italicised Ramsey's quote of Keynes)

The second interpretation of Carabelli's position seems to be the same as Ramsey's, that probability relations are somehow objective, but that there is no one probability relation which holds between any given set of propositions. This interpretation makes Keynes inconsistent. As shown in the discussion of Keynes's adoption of Russell's epistemology, Keynes held that direct knowledge was indubitable, and that probability relations were directly known. Keynes referred to propositions directly known as "self-evident", and either Keynes contradicted himself, or this interpretation of Keynes is incorrect. (In fact, this shows that Carabelli is incorrect when she states that Keynes did not believe that probability relations are self evident).

However, I believe that Ramsey (and, perhaps, Carabelli too) has misinterpreted Keynes. Ramsey seems to believe that there are only two possible positions: first, that the probability relation is always intuitable, and there can be no disagreement about what the relation is provided the correct conditions for intuiting prevail; second, that we can be wrong about the probability relation, and hence that we can never know the true probability relation between any set of propositions. I agree with Ramsey that if these are the only alternatives, then it would appear that Keynes did contradict himself, because he seems to adhere to both positions (and I would thus argue that Carabelli's interpretation of Keynes's position makes it inconsistent). But there is a third position, viz that only in some cases can we be certain about the correct probability relations, and this, I will now argue, is the position that Keynes took.

It does not seem to me that Keynes believed that our logical intuition is absolutely limited. Ramsey interprets Keynes's claims as being much stronger than they are: Ramsey seems to believe that for Keynes we can never intuit probability relations with certainty, and hence, that we could never determine the one correct probability relation between a given set of propositions. However, in the passages above, Keynes is discussing the limits of our reasoning in some cases, but not all. (O'Donnell 1991 interprets Carabelli as holding the second position, and makes the same point about it. This is also consistent with Moore's account of intuitions: Moore held that we had to employ a method to guard against confusion as to what we can intuit.) Keynes can reconcile the first and second positions by assuming that we can intuit logical truths (with certainty) in some, but not all, cases. Of course, Keynes would have to give a method of adjudicating between competing intuitions, and this seems impossible, as I will argue later. However, it seems that Ramsey's criticism (in this case) is misdirected. I further conclude that Keynes's account of justification did not differ significantly from Moore's.

4. The failure of intuitionism

We have seen that in his attempt to found the probability calculus on inductive logic, Keynes adopted Moore's account of justification. I have argued that Moore's account implies that the property "good" is essentially unfathomable. While it may be that ethical properties are inherently mysterious, I will argue in this section that there is no reason to think that probability is. I will now argue that these problems lead to the failure of Keynes's programme of providing foundations for the probability calculus (and in particular of justifying the Principle of Indifference), and thus demonstrate the inability of intuitionism to provide apriori foundations for methodology.

Keynes requires probability to be like goodness: a simple property known by intuition (which may or may not be non-natural). Hence, 'probable', like good, may apply to one, but not another, physically identical, situation; intuitionism fails to connect probability to conduct; finally, intuitionism makes debate impossible about probability, with implausible consequences. This leads to the conclusion that the putatively intuited relation does not seem to exist. However, Keynes's account of probability can be shown to be wrong: Keynes claimed to be able to intuit a version of the Principle of Indifference free of difficulties, and he was wrong. Hence the unreliability of (supposedly indubitable) intuitions can be demonstrated. I will take each of these points in turn.

I pointed out earlier that, although Keynes never explicitly stated that probability was non-natural, it seems likely that he thought it to be so. It seems that, if probability is a non-natural property, that is, independent of the physical properties of a state of affairs, then there is no reason to suppose that one state of affairs at a particular time and place will have the same probability as an identical state of affairs at another. Of course, it may be objected that there must be some difference, such as a difference in time or place. However, the point is that if all relevant parameters are the same, then intuitionists cannot be sure that the probabilities will be the same. If probabilities could be determined by making physical measurements, then probability would be natural, and thus ascertainable by means other than intuition.

Nor can Keynes give a motivation for obeying the probability calculus, if probability relations are non-natural. The non-naturalness of the probability relation implies that violation of the probability calculus would have no physical effect. If it did have such an effect, then once again, we would be able to detect probabilities by physical means. Thus it does not seem that Keynes can give a reason for assigning hypotheses numerical weights of evidence in accordance with the probability calculus. So if probability is "non-natural" then we may safely ignore it, and it can provide no motivation for a normative methodology. If, however, probability is "natural," then it seems that we could give a naturalistic account of it. If probability is a natural property, then even though it may not be definable in some ultimate sense, it may be possible to associate with it some other property which could be investigated by science. Just as the colour yellow can be identified with a certain range of frequencies of the electromagnetic spectrum, so too could probability be associated with, say, limiting relative frequency, or with actual degrees of belief, and so the study of probability relations would belong to the natural sciences. There may some subjective experience, a quale, associated with the intuition of a particular probability relation (a possibility which seems highly implausible), but it is difficult to see what would be missing from an account of probability which referred only to its physical aspects, and which ignored the associate quale.

Also, one may always claim (as Keynes did concerning the Principle of Indifference) to have an apriori intuition that a rule is correct, but this leaves one unable to reply to someone who reports some quite different intuition that another rule, inconsistent with the first, is apriori. At least one of the alleged intuitions must be false, but intuitionism does not provide a means for telling which one. Moreover, some people who have seriously reflected on the probability calculus report that they have no intuitions of the type described by Keynes, e.g., Ramsey:

> ...there really does not seem to be any such things as the probability relations he [Keynes] describes. He supposes that, at any rate in certain cases, they can be perceived; but speaking for myself I feel confident that this is not true. I do not perceive them, and if I am to be persuaded that they exist it must be by

> argument; moreover I shrewdly suspect that others do not perceive them either, because they are able to come to so very little agreement as to which of them relates any two given propositions. (Ramsey 1978, p. 62)

This raises a problem for Keynes: first, Keynes held Ramsey in high regard, as is evident from his laudatory obituary of Ramsey (Keynes 1933 [1972], pp. 335-346), and as is well documented from other sources (see, for example, Moggridge 1992, pp. 364-365). Second, Ramsey was a brilliant logician, and widely acknowledged as such. So Keynes could hardly respond that Ramsey lacked a power of logical intuition which others possessed. Indeed, Keynes never explained, and does not seem able to explain, Ramsey's professed lack of intuitive power. This problem seems more acute for Keynes than Moore, since we can pick out good logicians better than we can pick out good persons.

When discussing Moore's intuitionism, Keynes addressed this very problem of disagreements over intuitive judgements, and seemed to concede that it discredited the intuitionistic approach: Keynes pointed out that those who won arguments among adherents of Moore's intuitionism were those who employed the best rhetorical devices:

> How did we know what states of mind were good? This was a matter of direct inspection, of direct unanalysable intuition about which it was useless and impossible to argue. In that case who was right when there was a difference of opinion? There were two possible explanations. It might be that the two parties were not really talking about the same thing, that they were not bringing their intuitions to bear on precisely the same object... Or it might be that some people had an acuter sense of judgment, just as some people can judge a vintage port and others cannot. On the whole, so far as I remember, this explanation prevailed. In practice, victory was with those who could speak with the greatest appearance of clear, undoubting conviction and could best use the accents of infallibility. Moore at this time was a master of this method—greeting one's remarks with a gasp of incredulity—*Do* you *really* think *that*, an expression of face as if to hear such a thing said reduced him to a state of wonder verging on imbecility, with his mouth wide open and wagging his head in the negative so violently that his hair shook. *Oh!* he would say, goggling at you as if either you or he must be mad; and no reply was possible. (Keynes "My Early Beliefs" 1938 [1972], pp. 437-438)

As we noted above (section C), there are difficulties in interpreting the remarks on Moore in Keynes's "My Early Beliefs." Perhaps (as Moggridge 1992 suggests) Keynes was exaggerating Moore's behaviour for comic effect (the essay was written for a meeting of friends: an evocative account of the occasion can be found in Levy 1975). But clearly there is an element of truth in this account of disagreement between intuitionists, namely, that argument is impossible. This elimination of debate is a bad thing, since as we have seen, Keynes proposed a theory which led to contradictions. As with tastes, there seems to be no disputing intuitions, and this suggests that like taste, there is no objective truth about the matter of such intuitions.

Finally, as noted above, Keynes adopted a modified form of the Principle of Indifference which leads to contradictions. This seems to show that he was in fact wrong about his logical intuitions, and thus that logical intuitions are highly unreliable. There is also a debate as to whether Keynes changed his views on probability (adopting a subjectivist approach) under the influence of Ramsey (Moggridge 1992, p. 623 provides a survey and references). I do not take a position on this question, but note that if Keynes did change his views, it once again highlights the unreliability of intuitions.

The main concern of this chapter has been the plausibility of intuitionistic foundations for a non-natural probabilistic methodology. Keynes went the furthest in developing such foundations, taking his cue from G.E. Moore's intuitionistic foundations of ethics. As is well known, Moore's account is not satisfactory, and it is not surprising that Keynes's foundations for probability are also not. In fact, I have argued that Keynes's programme encounters more severe difficulties. The major difficulties that this programme faces are epistemological: even if a non-natural property did exist, it seems impossible that we could in fact know it. Moreover, there are numerous reasons to think that there are no such properties. In light of these problems, I conclude that Keynes's account of justification, and hence an intuitionistic account, is unsatisfactory.

There are two further approaches to non-naturalism which follow on from the failure. The first is the analytic method, an apriori approach which requires neither the epistemology nor the metaphysics of the intuitionistic approach (this is the subject of chapter 2). The second and, I shall argue, related, approach is to minimize the set of intuited properties necessary for the probability calculus (this is the subject of chapters 3). I will argue in the next two chapters that these problems are endemic to non-naturalist accounts of justification of methodologies.

CHAPTER 2: The Apriori as the Analytic: Carnap's Logical Foundations of Probability

> But what, then, is left over for *philosophy*, if all statements whatever that assert something are of an empirical nature and belong to factual science? What remains is not statements, nor a theory, nor a system, but only a *method*: the method of logical analysis. (Carnap, 1932 [1959], p. 77)

The next attempt I consider to found an inductive methodology on apriori grounds alone is the attempt to demonstrate that the probability calculus is analytically true (or analytic, for short). I take Rudolf Carnap as the most important advocate of this approach. Carnap's views were developed in the milieu of logical positivism. This movement grew out of the Vienna Circle, which rejected metaphysics as meaningless. Members of the Circle agreed with Hume that all meaningful statements are either logical or empirical, corresponding, respectively, to the apriori and the aposteriori. (They also rejected Kant's claim as to the existence of synthetic apriori truths, hence equating the aposteriori with the synthetic, and the apriori with the analytic). Accordingly, they stressed analyticity as the basis of all apriori knowledge (a logical statement is said to be analytic if it is true solely in virtue of its meaning). According the Vienna Circle, the purpose of philosophy was divide the true statements of science into these categories (see the motto of this chapter, but also: "*Philosophy is the logic of science*, i.e., the logical analysis of the propositions, proofs, theories of science..." Carnap 1934 [1967], pp. 54-55.) Carnap attempted to provide a non-empirical (hence apriori, analytic) foundation for an inductive probabilistic logic by showing that the axioms of inductive probability are analytic: "...all principles and theorems of inductive logic are analytic." (Carnap 1950, p. v.)

I shall examine in this chapter how Carnap attempted to provide analytic foundations for an inductive methodology. To do this, however, I must explicate Carnap's views on justification, which are, unfortunately, not always clear. Although Carnap is usually interpreted as providing a non-empirical justification (for example, Braithwaite referred to Keynes and Carnap as "...adherents of a non-empirical theory of probability who attempted to produce a formal logic of induction..." (Braithwaite 1968, p. 196)), and, indeed, Carnap himself claimed to provide such an account (see above, and also below, section B.3), Carnap also sometimes appears to have expounded conventionalist, and at other times empirical, foundations of methodology. Exegesis is required to sort out these three incompatible strands. I shall conclude that, although Carnap began from a conventionalist theory of justification, he mistakenly used this theory to argue for an apriorist theory of justification. At the same time, Carnap also attempted to provide an empirical theory of justification. I conclude that there are good reasons for Carnap's changing views: I shall that argue Carnap failed in his attempt to provide analytic foundations for the probability calculus and indeed, that all such attempts are doomed to failure.

A. Frameworks, Internal Questions, and The Method of Logical Analysis

Carnap rejected philosophical methods that depend on supposed intuitions in favour of a method of philosophical analysis, that is, of identifying analytically true relations. For Carnap, identification of analytically true statements was dependent on the language in which the analysis was undertaken. Languages, however, are not the sorts of thing which can be right or wrong. For Carnap this implied that the adoption of a language is dependent on conventional considerations. Carnap's conventionalism with respect to the choice of a language is made clear by his distinction between "internal" and "external" questions.

Carnap held that there was "a fundamental distinction between two kinds of questions concerning the existence or reality of entities." (Carnap 1950 [1967], p. 73) To elucidate this distinction, Carnap introduced the notion of a framework: "If someone wishes to speak in his language about a new kind of entities, he has to introduce a system of new ways of speaking, subject to new rules; we shall call this procedure the construction of a linguistic *framework* for the new entities in question." We can then distinguish between internal questions, which are "questions of the existence of certain entities of the new kind *within the framework*" and external "questions concerning the existence or reality *of the system of entities as a whole...*" Internal questions may be answered "either by purely logical methods or by empirical methods, depending upon whether the framework is a logical or a factual one." As an example Carnap considered what he took to be our ordinary language, the "thing language." Using this language allows one to answer, by empirical investigation, the internal question "Are unicorns and centaurs real or merely imaginary?" But Carnap considered the external questions like "Is there a thing world?" to be "framed in a wrong way" (or "pseudo-questions," "non-cognitive." (ibid, p. 75)). The only way to make sense of such a question, according to Carnap, is to construe it as a "practical question, a matter of a practical decision concerning the structure of our language." (ibid, p. 73) The decision to adopt a language is thus "practical" and not "theoretical": it is determined by the consequences of its choice:

> To be sure, we have to face at this point an important question; but it is a practical, not a theoretical question; it is the question of whether or not to accept the new linguistic forms. The acceptance cannot be judged as being either true

> or false because it is not an assertion. It can only be judged as being more or less expedient, fruitful, conducive to the aim for which the language is intended. (ibid, p. 78-79)

The above might lead one to believe that Carnap was a thoroughgoing conventionalist. This is not correct, however, since he thought the choice of a language could be constrained by empirical considerations. Carnap held that while the initial choice of a framework is a conventional, its continued acceptance is governed by empirical considerations:

> The acceptance or rejection of abstract linguistic forms, just as the acceptance or rejection of any other linguistic forms in any branch of science, will finally be decided by their efficiency as instruments, the ratio of the results achieved to the amount and complexity of the efforts required. (ibid, p. 83)

> ...the work in the field will sooner or later lead to the elimination of those forms which have no useful function. (ibid, p. 83-84)

Carnap summed up his views on the adoption of a language in his famous Principle of Tolerance: "*Let us be cautious in making assertions and critical in examining them, but tolerant in permitting linguistic forms.*" (ibid, p. 84) We might call Carnap's view conventionalism-cum-empiricism. (The interpretation I offer seems to be the same as given by Richard Creath 1990, who also argues that for Carnap, analytic, apriori truths are determined by convention.)

Carnap's views led him to reject intuitionism, both in general and specifically with reference to theories of probability. Generally, Carnap held that it is not possible to draw a distinction between "data (that which is immediately given in consciousness, e.g., sense-data, immediately past experiences, etc. [knowledge by acquaintance]) and the constructs based on the data [knowledge by description]." (1950 [1967], p. 82). This distinction, as shown in Chapter 1, was foundational for Moore's and Keynes's intuitionism. And, indeed, Carnap specifically rejected intuitionism: "...the semanticist does not in the least assert or imply that the abstract entities to which he refers can be experienced as immediately given either by sensation or by a kind of rational intuition." Thus Carnap explicitly rejected the idea that intuitions can be invoked as evidence for the existence of an abstract entity. (Further, Carnap saw intuitionism as bad psychology: "An assertion of this kind [that we intuit abstract entities] would indeed be very dubious psychology." (ibid, p. 83).)

Carnap specifically rejected intuitionism as a means of justifying a probabilistic inductive logic. We find this rejection in Carnap's discussion of Koopman's view that probability values are given by intuition:

> The results of this special kind of intuition seem to be regarded as not subject to rational examination (except for questions of consistency) and therefore not capable of rational reconstruction. This view is similar to, and probably influenced by, Keynes's conception of probability as undefinable and based on in-

> tuition. In contrast to Koopman's view, I am convinced that it is possible to give a rational reconstruction or explication for the comparative concept of confirmation, and I believe, moreover, that it is possible to define an explicatum without using any other terms than those of deductive logic, hence, of semantics." (Carnap 1950, p. 453)

Carnap held that inductive logic was to be obtained from the analysis of a given language. As we shall see, this language was provided by science. The job of the philosopher was to 'explicate' or 'rationally reconstruct' this logic by means of analysis (how this was to be done will be examined in the next section). However, Carnap argued that what he was explicating was an 'objective' notion of probability, and his interpretation of 'objective' misled him into arguing for apriori foundations for probability. The next section details his changing views.

B. Carnap on the Justification of an Inductive Logic.

In this section I survey Carnap's attempts to justify the probability calculus with his conventionalist, then his empirical, and finally his apriori accounts of justification. I shall show that Carnap in the end settled on the apriori account, perhaps because of the difficulties raised by the other two. I defer a detailed account of these difficulties to Parts II and III, but in the following sections I shall argue that any attempt to found an inductive methodology in apriori fashion using logical analysis is bound to fail.

1. Carnap's Conventionalism

Carnap held that analytic truth is relative to a conventionally chosen linguistic framework. He also believed inductive logic to be analytic. It would thus follow that, from a Carnapian point of view, inductive logic is conventional. Indeed, in §53 of the *Logical Foundations* Carnap notes some "conventions" (for example, that the probability of a tautology conditional on any evidence is 1) which "practically all authors who use probability$_1$ [logical probability] as a quantitative concept... have accepted..." (1950, p. 284) Carnap states these conventions as definitions, not as axioms: "Some authors have laid down these conditions, or similar ones from which these follow, as axioms. We shall not do so; our system of inductive logic will not be based on axioms but only on explicit definitions." (Carnap, 1950, p. 284.) Carnap is referring to the difference between explicit and implicit definitions: implicit definitions are uninterpreted axioms, whereas explicit definitions reduce the concepts to be defined to other concepts. Carnap presumably wished to reduce the notion of probability as commonly used by scientists through analysis to purely logical notions: indeed this seems to be what he meant by 'rational reconstruction.'

The definitions of probability were to be obtained by "rational reconstruction", which Carnap defined as "a more exact formulation... of a body of generally accepted but more or less vague beliefs." (1945, p. 95) As Carnap used the term, "rational reconstruction" is interchangeable with "explication." An explication is an attempt to make a term in ordinary language, the explicandum, precise. The product of an explication is an explicatum: "There is a certain term ('confirming evidence', 'degree of confirmation', 'probability') which is used in everyday language and by scientists without being exactly defined, and we try to make the use of these terms more precise or, as we shall say, to give an *explication* for them." (Carnap 1950, p. 2) Carnap noted that what scientists do and what they say are often two different things: "Those who work in the history of science or the methodology of science are familiar with the fact that there is frequently a discrepancy between what an author actually does and what he says he does; in particular, between the sense in which he actually uses a term or a sentence and the sense which he explicitly attributes to it. This holds especially for abstract terms and general principles." (Carnap 1950, p. 37). Carnap thus recommended a study of scientists' actions as well as their words: "Consequently, in order to find out which sense a certain term has for the author, it is often not sufficient to look at his explicit explanations. We should also examine how he uses the term and, especially, how he argues pro or con statements in which the term occurs." (1950, p. 37)

However, according to Carnap, not any explication will do. He proposed that an explicatum was satisfactory if it met four requirements: First, "The explicatum is to be *similar to the explicandum* in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used..."; second, "The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts"; third, "The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept)"; and fourth, "The explicatum should be as *simple* as possible..." (Carnap 1950, p. 7) It seems however, that the second and the third were the most important criteria for Carnap, since he held "close similarity is not required, and considerable differences are permitted" and that the explicatum should be "as simple as the more important requirements... [similarity, exactness and fruitfulness] permit." (Carnap 1950, p. 7).

Unfortunately, Carnap left these notions vague, and so did not give any clear guide to what a successful rational reconstruction might look like. This problem is compounded by the fact that he held that there is no one correct explication of probability, since the meaning of probability in ordinary language includes both probability$_1$ (the logical conception, to be discussed in detail below) and probability$_2$ (the frequentist conception of probability as limiting relative frequency). In fact, Carnap saw this as a general feature of all attempts at explication:

> In a problem of explication the datum, viz., the explicandum, is not given in exact terms; if it were, no explication would be necessary. Since the datum is inexact, the problem itself is not stated in exact terms; and yet we are asked to give an exact solution. This is one of the puzzling peculiarities of explication. It follows that, if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it is right or wrong. Strictly speaking, the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer. (Carnap 1950, p. 3-4)

Hence, it would seem that the choice of one explication over another is pretty much a matter of pure convention. Since the empirical constraints on an explication are vague, and since more than one explication may meet these constraints, it seems that there will be a number of equally acceptable explications.

However, the term "rational" in "rational reconstruction" seems to imply an element of justification. Carnap denied this implication, differentiating between two types of justification. The first is a weaker form: a rational reconstruction is justified inasmuch as it captures adequately our prescientific concepts. According to Carnap, c*, one of the inductive methods he studied in detail, and which I shall discuss below, is justified as a rational reconstruction in this sense. But the second sense of justification is "to show the validity of the new theory and thereby of the given beliefs," or in terms of an inductive logic, to demonstrate "the validity of our or any other proposed system of inductive logic..." According to Carnap, "This is the genuinely philosophical problem of induction," (1945, p. 96) which he regarded as insoluble. So, it is not possible to show that the probability measure chosen as a rational reconstruction is the one true inductive measure: "The situation just described has sometimes been characterized by saying that a theoretical justification of induction is not possible, and hence, that there is no problem of induction. However, it would be better to say merely that a justification in the old sense is not possible." (1945, p. 96) Hence, it would seem that Carnap was left with no choice but conventionalism, since he felt only the first type of justification is possible: while our reconstruction may be justified inasmuch as it explicates ordinary inductive practices, no "external" justification of these practices is to be had. So, it would seem that Carnap would agree that many inductive logics may be obtained from the process of rational reconstruction ~~useful~~.

Unfortunately, the matter is not so clear, since Carnap it was the task of philosophers to explicate an objective theory of probability since the term 'probability', as used by scientists, is an "objective" notion, as all previous writers on the logical conception of probability (including Laplace, Keynes, Jeffreys, Ramsey and Bernoulli) had assumed:

> It seems to me that, on the basis of the discussions of this section, it is plausible to assume that for most, perhaps for practically all, of those authors on probability who do not accept a frequency conception the following holds. (i) Their theories of probability are objectivistic... (ii) The objective concept which they mean, clearly or vaguely, as their explicandum is something similar to probability$_1$; in the classical period the explicandum is often not yet quite clear; but it seems that in the course of the historical development the concept of probability$_1$ emerges more and more clearly. (1950, p. 50-51)

Since "probability$_1$" was a logical conception, Carnap followed what he took to be logicians' conceptions of "objectivity." 'Objective' in this sense means that, as we shall see, statements have a unique probability, independent of person and time. In other words, Carnap rules out the subjectivist account.

According to Carnap, "practically all logicians" agree with the "objectivist conception of logic." (1950, p. 39), and their working practices demonstrate this: "[I]n order to find out whether a certain conclusion follows from given premises, they [logicians] do not in fact make psychological experiments about the thinking habits of people but rather analyze the given sentences and show their conceptual relations." (1950, p. 39) 'Objective' does not mean 'apriori independent of agreed upon conventions.' In accordance with the above discussion of Carnap's notion of explication, objective must mean correct within a given logical system, the logical system being chosen by extralogical considerations. Unfortunately, Carnap seemed to forget his earlier account of explication, and equated 'objective' with something like 'correct in all logical systems.'

This shift can be found in Carnap's discussion of logical consequence. As an example he takes contraposition. Taking *i* to be "If A then B", *j* to be "If not B then not A" then, in first-order logic, *i* implies *j*. According to Carnap, *j*'s being a logical consequence of *i* is independent of anyone's beliefs; and in this sense the relation of logical consequence is objective: "The relations are objective, not subjective, in this sense: whether one of these relations does or does not hold in a concrete case is not dependent upon whether or what any person may happen to imagine, think, believe or know about these sentences..." (Carnap 1950, p. 38) This is surely correct, given the framework of first-order logic. However, Carnap argues that 'objective' means that either contraposition is a correct rule of inference, or it is not, regardless of the logical system:

> Suppose that a person X believes at the present time that *j* is a logical consequence of *i*, while at an earlier time he believed that this was not the case. That the relation is objective is meant in this sense: the change in X's belief about the

> relation has no effect upon the status of the relation itself; if his present belief is right (as I think it is), then his former belief was wrong; and, if his former belief was right, his present belief is wrong. (1950, p. 38)

While Carnap leaves open the possibility that contraposition is not a correct rule of inference for conditionals (and indeed, it is not in Lewis-Stalnaker counterfactual conditional logics), he is clear that it either is or is not apriori correct. Carnap has shifted from 'objective within a logic' to 'objective apriori,' in contradiction to his account of explanation.

Further evidence for this shift can be found in Carnap's discussion of "psychologism." He characterises "A primitive psychologistic explanation of the relation of logical consequence" as "That *j* is a logical consequence of *i* means that, if somebody believes in *i*, he cannot help believing also in *j*." According to Carnap, no one would ever adopt such a form of psychologism, "because its inadequacy is too obvious," since "Taken literally, the explanation given would require us to investigate the statistical results of series of psychological experiments." Carnap concludes from this that "There are not many logicians who would regard this procedure as appropriate." (1950, p. 41) This is undoubtedly correct, and so Carnap is right to reject such a notion of logic for explication. However, Carnap goes further, and rejects psychologism as a possible foundation for logic. For example, he quotes approvingly a parody of a psychologistic procedure:

> A nice illustration, though not meant quite seriously, of primitive psychologism in arithmetic—which is part of deductive logic—is the following passage by P.E.B. Jourdain (*The philosophy of Mr. B*rtr*nd R*ss*ll...*): "I sometimes feel inclined to apply the historical method to the multiplication table. I should make a statistical inquiry among school children, before their pristine wisdom has been biased by teachers. I should put down their answers as to what 6 times 9 amounts to, I should work out the average of their answers to six places of decimals, and should then decide that, at the present stage of human development, this average is the value of 6 times 9." (Carnap 1950, p. 41)

And, in the glossary to his *Logical Foundations*, Carnap defines psychologism as the "*wrong* interpretation of logical problems in psychological terms." (Carnap 1950, p. 581, italics added) This claim of wrongness is in contradiction with his notion of explication. If all logicians decided to adopt such a basis, then, by Carnap's theory of explication, it would be the job of the philosopher to explicate this view. However, Carnap never seems to have noticed this conflict. Instead, he extended his views to inductive logic.

As I shall show in section C, Carnap held that inductive logic was entirely definable in terms of deductive logic, and that the relations of deductive logic hold irrespective of the psychological state of anyone considering these relations, he also held that it would be a mistake to base an inductive calculus on how people actually make inductive judgements: "a definition of an explicatum for probability$_1$ must not refer to any person and his beliefs but only to the two sentences and their logical properties *within a given language system.*" (1950, p. 43, italics mine) Carnap did not argue against the subjectivist view. He simply dismissed it out of hand, asserting that such an approach could not be useful for science:

> It cannot, of course, be denied that there is also a subjective, psychological concept for which the term 'probability' may be used and sometimes is used. This is the concept of the degree of actual, as distinguished from rational, belief: 'the person X at the time *t* believes in *h* to the degree *r*.' This concept is of importance for the theory of human behavior, hence for psychology, sociology, economics, etc. But it cannot serve as a basis for inductive logic or a calculus of probability applicable as a general tool of science. (1950, p. 51)

Neither here, nor elsewhere in *Logical Foundations* does Carnap offer arguments for this conclusion.

Carnap's confusion over the explication of an objective concept no doubt helps to explain his later turn to apriorism. This apriorism is inconsistent, however, with his earlier conventionalist account of explication. But before he had adopted a fully aprioristic approach, Carnap also advocated a modified conventionalist approach: he attempted to justify methods on the basis of empirical success.

2. Empirical Foundations

Pure conventionalism would presumably allow a purely pragmatic choice among the indenumerably many inductive methods. As I shall show in section C, Carnap sometimes sought to narrow the range of acceptable methods by the criterion of potential success in giving accurate forecasts, and rejected the conventionalism implied by his notion of explication: "the decisive justification of an inductive procedure does not consist in its plausibility, i.e., its accordance with customary ways of inductive reasoning, but must refer to its success in some sense." (Carnap 1945, p. 96) Carnap defined the success of an inductive method as the degree to which it would lead to a match between probabilities assigned to an event and observed relative frequencies of that event. Of course, in the short run, as Carnap notes, infinitely many inductive methods will do this. Therefore he proposed to investigate means for determining which inductive methods will be more successful in the long run: "...we need a more general and stronger method for examining and comparing any two given rules of induction in order to find out which of them has more chance of success." (ibid, p. 97) However, Carnap, in 1945, had not developed a means for comparing inductive measures: "much further investigation will be needed."

Carnap found this means of classification later: a single real-valued parameter λ, which I shall describe in section C. In the preface to his 1950, Carnap claimed that this classification could be used to characterise the inductive success of a method in any possible world (Carnap represented possible worlds using state descriptions, as I shall show below). He acknowledged that since we do not know which possible world we actually inhabit, and since no one inductive method is superior to every other method, we do not know which would be the best method to employ: "[the] total structure [of the actual world] is not known to us... therefore [we] cannot know which of... two inductive methods would be more successful in the long run." However, he argued that if one method were more successful than another method in a large enough set of worlds, we might be inclined to choose it:

> Suppose, for example, that in comparing two given inductive methods we find that the number of those state-descriptions in which the second method is more successful is a million times as large as the number of those in which the first method is more successful. Then it may well be that this result would influence us against regarding the first method as more adequate than the second and against choosing the first in preference to the second for determining our practical decisions in the actual world... (Carnap 1950, p. x)

Thus Carnap gives an empirical reason for choosing an inductive method based on a view of gambles: you might be more likely to have a more successful inductive method if you choose an inductive method which is successful in a large enough class of possible worlds (described, as I shall show, by what Carnap termed "state-descriptions"). The problem with this argument, as Carnap pointed out, is that we do not know which world we inhabit. But we may believe that we inhabit a world in which a method which is overall less successful (however this might be determined) is in fact the correct method. We would thus not be inclined to accept the method which is more successful in a larger number of worlds.

It could be that Carnap was giving a practical argument based on the psychological fact that we are more inclined to choose methods which might be successful in this manner. Carnap claimed that his characterisation of inductive logic would allow the construction of inductive methods which would be maximally successful in a given world: "it is easily possible... to construct that particular inductive method which is most successful for the given world." (Carnap 1950, p. xi) The only reason one would wish to do this, it seems, is that one would wish to choose the most successful inductive method for the world in which we happen to believe that we inhabit: this is an empirical justification for an inductive method, since we would choose a method aposteriori, based on our beliefs of what the world is like.

Carnap never did publish his planned second volume of the *Logical Foundations*. He did, however detail his method for characterising inductive methods in 1952, in his *The*

Continuum of Inductive Methods. In this work, Carnap noted that numerous deductive logics have been developed (such as intuitionistic logic and three-valued logic), and that the situation was the same in inductive logic, that is, that there are also numerous possible inductive logics. But he also claimed that further advances might narrow the choice among both inductive and deductive logics: "The fact that there are at the present time irreconcilable differences may be merely a matter of historical contingency due to the present lack of knowledge in the field of inductive logic. If so, it would be conceivable that at some future time, on the basis of deeper insight, all will agree that a certain inductive method is the only valid one." Nonetheless, Carnap also acknowledged that there may be more than one inductive logic: "it may be that the multiplicity of mutually incompatible methods is an essential characteristic of inductive logic, so that it would be meaningless to talk of 'the one valid method.'" (Carnap, 1952, p. 7) Carnap then seemed to suggest that a choice of an inductive logic could be made on other, non-apriori, considerations:

> This does not necessarily imply that the choice of an inductive method is merely a matter of whim. It may still be possible to judge inductive methods as being more or less adequate. However, questions of this kind would then not be purely theoretical but rather of a pragmatic nature. A method would here be judged as a more or less suitable instrument for a certain purpose. I shall not try to discuss this problem here, still less to solve it; but I may indicate that at the present time I am more inclined to think in the direction of the second answer. (ibid)

In fact, these other considerations were empirical and pragmatic: in choosing an inductive method Carnap suggested that we take into account how easy the system is to use, aesthetic considerations, and how closely the method corresponds to observed relative frequencies:

> [An individual choosing an inductive method] will take into consideration the performance of a method, that is, the values it supplies and their relation to later empirical results, e.g., the truth-frequency of predictions and the error of estimates; further, the economy in use, measured by the simplicity of the calculations required; maybe also aesthetic features, like the logical elegance of the definitions and rules involved. (ibid, p. 55)

This justification of an empirical method is, for the reasons stated above, also unsuccessful. However, I shall defer detailed discussion of such empirical justifications to Chapter 4. We may conclude, however, that Carnap did indeed hold at least three different views of justification of the probability calculus. In 1945 he leaned towards an empirical view, in 1950 he leaned towards apriorism, and in 1952 he leaned once again to the empirical view, and at the same time seems to have accepted a conventional basis for methodology.

3. Carnap's apriorism

Carnap's turn to apriorism, as we have seen, began in 1950. Also, Carnap included in his

Logical Foundations an appendix on "Inductive Logic as a Rational Reconstruction," which seems to have been based on §16 of his 1945. Carnap repeated several sentences from his 1945 about c^* being intended as a rational reconstruction, but he does not despair of showing that one inductive method is valid. Instead he repeats the claim that c^* is a valid reconstruction of our ordinary inductive beliefs, and then claims that a rational reconstruction aims for "results which are more systematized, more consistent, and in certain points more correct than customary ways of thinking." (1950, p. 576) Carnap defined one inductive method "as more correct or more reasonable than another if it is in better accord with the basic principle of inductive reasoning." (ibid) This basic principle "says that expectations for the future should be guided by the experiences of the past. More specifically: what has been observed more frequently should, under otherwise equal conditions, be regarded as more probable for the future." (ibid) And, according to Carnap, "in the points where there is a divergence [between "customary inductive thinking" and c^*], the theory of c^* is more correct than customary thinking..." (ibid, p. 577)

Further, in his 1962 (a condensed version of his 1971), Carnap seems to have adopted an extra-linguistic, apriori basis for inductive logic. In this paper, Carnap rejected a "psychological" interpretation of probability in favour of a "rational or normative decision theory" based on "requirements of rationality." (1962, p. 303) Carnap based his inductive logic on axioms, as opposed to definitions, and claims that they are "justified by general considerations of rationality." (ibid, p. 307) Unfortunately, the only defense Carnap offered for his axioms is that disagreement with them is "obviously unreasonable." (ibid) However, Carnap did acknowledge that his axioms did not pick out a unique inductive measure, and remained agnostic as to whether or not one would ever be found. (ibid, p. 316)

Carnap also referred to intuition to justify a choice of axioms for inductive logic. In his 1963 (p. 978) he summarised his view of justification in terms of coherence with intuitions, and characterised justification as apriori:

> It seems to me that the reasons to be given for accepting any axiom of inductive logic have the following characteristic features...
> (a) The reasons are based upon our intuitive judgments concerning inductive validity, i.e., concerning inductive rationality of practical decisions (e.g., about bets)
> Therefore:
> (b) It is impossible to give a purely deductive justification of induction.
> (c) The reasons are a priori.

Finally, in his 1968, Carnap seems to reject his earlier attempts at an empirical justification of an inductive method in favour of an intuitionistic one. He summarises this attempt, and then says that it is not incorrect:

> I said [in *The Continuum of Inductive Methods*] that a C-function [a probability measure] used for making decisions is a kind of a tool. If we are not satisfied with the working of a tool, we change it. If our experiences seem to indicate that our tool, the C-function, does not work well, we might change it on the basis of those experiences. Today I would not say it is wrong to proceed in this way. If you want to, you may use past experience as a guide in choosing and changing your C-function. (Carnap 1968, p. 264)

But in the next sentence, Carnap asserts that a C-function is apriori: "But in fact it is never necessary to refer to experiences in order to judge the rationality of a C-function." Carnap then draws an analogy with arithmetic. We may count objects to show that arithmetic actually works, but "most of us would agree that this is not necessary..." because "arithmetic is a field of a priori knowledge." (ibid, pp. 264-265) Carnap then asserts that "the same holds for inductive logic." Hence Carnap seems to believe that checking an inductive measure's success does not contribute to the justification of that inductive measure: the aposteriori success of a method is explained by the apriori justification of that method. In the discussion of his 1968, Carnap once again affirmed apriorism: "Today I am inclined to base the choice of an inductive method... not on observed facts, but only on a priori considerations." (ibid, p. 314)

Carnap also refers to "intuition" as a way of justifying an inductive method. Although he is quick to point out that intuition is not infallible, infallibility seems to be required: "If a person were unable to distinguish valid from invalid steps in inductive reasoning, even in the simplest cases, in other words, if he were inductively blind, then it would be hopeless to try to convince him of anything in inductive logic." (ibid, p. 265)

However, the point of the above exegesis is not merely to demonstrate that Carnap changed his mind, but also to show that Carnap can be interpreted in an apriorist, and can be criticised as such. Therefore, in the next sections, I will examine Carnap's attempt to construct an inductive logic, and show that this logic is not apriori justified. In the final chapter of this thesis, I will argue that Carnap's earlier, conventionalist, views on justification were superior.

C. The Logical Foundations of Probability

To see why it was necessary for Carnap, if he did adopt an apriorist view, to constrain the probability measures, it is useful to look at his account of deductive logic. According to Carnap, deductive relations were analytic because whatever is contained in the conclusion of a valid inductive argument is also contained in the premises. This containment could be determined solely by the meanings of the terms (how this was to be done will be described below), and illustrated in the box on the left in figure 3.

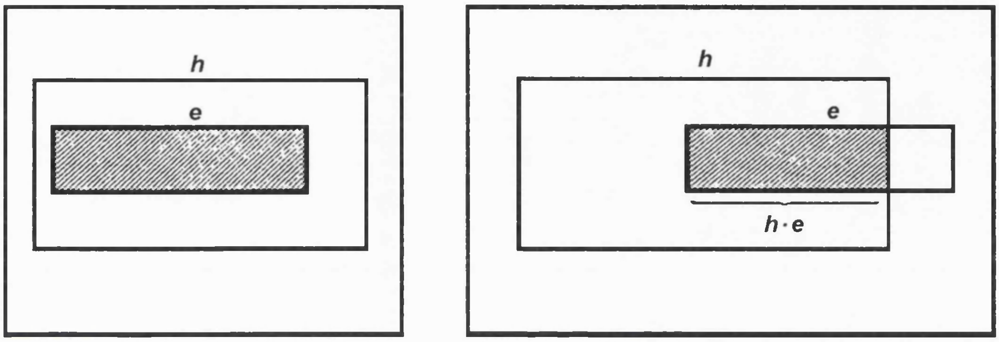

Figure 3: Deductive and Inductive Inclusion
(From Carnap 1950, p. 297)

Carnap wanted a similar notion of containment for probabilistic inductive logic, hence the box on the right of figure 3. The relation of partial entailment cannot be analytic if two persons who agree on the meanings of two statements A and B disagree as to the value of P(A | B).

For Carnap, analytic statements are statements that are logically true (Carnap 1945, pp. 73-74). Carnap defined logical truth in terms of state descriptions which are conjunctions of all atomic sentences or their negations. The simplest state descriptions are those of the primitive language $\mathcal{L}(A,n)$, that is, a language with a single monadic predicate A, n distinct individuals $a_1,\ldots a_n$, and the sufficient set of connectives $\neg$ and $\wedge$ (the others being defined as usual). We can then represent all possible states of affairs describable in $\mathcal{L}$ by means of the 2^n conjunctive sentences $\pm A(a_1)\wedge\pm A(a_2)\wedge\ldots\wedge\pm A(a_n)$ (where $+A(a_i)$ is $A(a_i)$ and $-A(a_i)$ is $\neg A(a_i)$). Each of these sentences is a state-description. Relative to a language, a state-description provides a non-decomposable partition of descriptions of states of affairs (in other words, a state description describes a possible world). Carnap defined the *range* of a statement as the set of state descriptions that are compatible with it.

Deductive logic can be described in terms of ranges: a sentence is a tautology if and only if it holds in all possible worlds (that is, it is compatible with the truth of every state description—i.e., it is logically true or analytic; it is a contradiction if and only if it does not hold under any state-description (that is, it is logically false and incompatible with all state descriptions); if it is neither a tautology nor a contradiction it is "factual (synthetic, contingent)"; h implies e if and only if the range of e is included in h (that is all state descriptions compatible with h are also compatible with e); finally, two statements are equivalent if and only if they have the same range.

Carnap interpreted inductive logic as a probability measure of ranges. This can be seen by defining a function m such that

1. $\sum_i \mathrm{m}(\mathrm{P}_i) = 1$, where the $\{\mathrm{P}_i\}_i$ is the set of state descriptions.
2. If h is not logically false, $\mathrm{m}(h)$ is equal to $\sum_j \mathrm{m}(\mathrm{P}_j)$, where $\{\mathrm{P}_j\}_j$ is the set of state-descriptions which make h true (i.e., $\{\mathrm{P}_j\}_j$ is the range of h).
3. If h is logically false, $\mathrm{m}(h) = 0$.

and a function c such that

4. $\mathrm{c}(h|e) = \mathrm{m}(e \wedge h)/\mathrm{m}(e)$.

Clearly, m is a probability measure, and c is a conditional probability function. $\mathrm{c}(h|e)$ is a normalised measure of the inclusion of the range of e in the range of h, and so is a measure of partial entailment. In virtue of the definition of the measure, any theorem which follows from the above three definitions will be analytic by Carnap's standard.

However, the statement that $\mathrm{c}(h|e) = r$ for a given h and e and a unique r is not analytic, since this is dependent on the values the function m assigns, and (almost) any value between 0 and 1 will satisfy the definitions. The axioms of the probability calculus alone do not uniquely determine particular assignments. Hence Carnap sought additional restrictions on probability measures in order to make statements of conditional probability analytic. Rules governing the assignment of particular probability values within the constraints of 1, 2 and 3 are called by Carnap inductive methods. However, Carnap's investigations led him to the discovery of infinitely many inductive methods, the λ-continuum. From this continuum of inductive methods, he concentrated on two measures, m^* and $\mathrm{m}^\dagger$. In the next section I describe these measures, and argue that Carnap's attempt to justify a particular inductive method failed.

D. The Continuum of Inductive Methods

In his *Logical Foundations* Carnap considered only two conditional probability measures, $\mathrm{c}^\dagger$ and c^*. Both are obtained by imposing symmetry conditions: the first by assigning state descriptions equal probability, and the second by assigning structure descriptions equal probability (I shall describe structure descriptions below). In this section I shall show how Carnap's investigations led him to indenumerably many inductive measures, each corresponding to a (real) value of a parameter λ. I shall also argue that Carnap's attempt to justify a particular measure failed.

$\mathrm{c}^\dagger$ is the conditional probability measure obtained by applying the Principle of Indifference to state descriptions. Since in $\mathscr{L}(\mathrm{A},n)$ there are 2^n state descriptions, if we were, using the Principle of Indifference, to assign equal probability to each of the state descriptions, they would each have probability $(1/2)^n$. This measure turns out to be insensitive both to sample size and to information about the proportion of a_is possessing A. This is proved thus: any sentence h in the language $\mathscr{L}(\mathrm{A},n)$ can be represented as a disjunction of q of the state descriptions ($0 \leq q \leq 2^n$), and thus $\mathrm{m}^\dagger(h) = q/2^n$. Consider the sentence $\mathrm{e}(r,k)$, which states that r of the k individuals $a_1,\ldots,a_k$ in a sample have A. This sentence is equivalent to a disjunction of state descriptions. A state description which describes r of k individuals having A can be constructed first by considering all the ways in which the property could be distributed among the individuals. This number is ${}^k\mathrm{C}_r$. However, this determines only the first k of the places in the state description. The state description can be 'completed' in 2^{n-k} ways, and hence the total number of state descriptions describing the sample is ${}^k\mathrm{C}_r\, 2^{n-k}$, and so the ($\mathrm{m}^\dagger$) probability of r individuals in a sample of k is ${}^k\mathrm{C}_r\, 2^{n-k}\, 2^{-n}$. The probability of the $(k+1)$th individual having A and r individuals from k having A, $\mathrm{A}(a_{k+1}) \wedge \mathrm{e}(k,r)$, is ${}^k\mathrm{C}_r\ 2^{n-k-1}\ 2^{-n}$, since one more state description is fixed. Hence,

$$\mathrm{c}^\dagger(\mathrm{A}(a_{k+1})|\mathrm{e}(k,r)) = \frac{\mathrm{m}^\dagger(\mathrm{A}(a_{k+1})\wedge\mathrm{e}(r,k))}{\mathrm{m}^\dagger(\mathrm{e}(r,k))} = \frac{2^{n-k-1}2^{-n}\,{}^k\mathrm{C}_r}{2^{n-k}2^{-n}\,{}^k\mathrm{C}_r} = \frac{1}{2}.$$

This equality holds regardless of the values of k and r, which means that the measure is insensitive to sample size—no matter how many individuals are examined, the probabilities will never change. Such a measure cannot be inductive, since it does not provide for learning from experience, and for this reason, Carnap rejected it. (1952, p. 38)

The measure m^* is based on equiprobability of structure descriptions. A structure description is a description of all the ways the property can be distributed among the individuals: this is done by forming a disjunction from all structure descriptions with the same number of predicates. For example, in the simple case of $\mathscr{L}(\mathrm{A}, 2)$, there are three structure descriptions: $\mathrm{A(a_1)\wedge A(a_2)}$, $(\mathrm{A(a_1)}\wedge\neg\mathrm{A(a_2)})\vee(\neg\mathrm{A(a_1)}\wedge\mathrm{A(a_2)})$, $(\neg\mathrm{A(a_1)}\wedge\neg\mathrm{A(a_2)})$. Assigning equal probability to all the structure descriptions is equivalent to the Laplacian Principle of Indifference (as opposed to the Keynesian Principle of Indifference, $\mathrm{m}^\dagger$), and so Carnap was able to derive the Rule of Succession, that is, a rule, first given by Laplace, that the probability of A occurring the $k+1$th time given it have already happened r times out of k is equal to $(r+1)/(k+2)$. Since this measure is sensitive to the samples size, Carnap held that it could serve as an inductive measure.

The measures $\mathrm{m}^\dagger$ and m^* Carnap discussed in *Logical Foundations* are only two among indenumerably many possible measures, and in *The Continuum of Inductive Methods* he developed a means to classify a large class of them. Carnap generalised the family of probability measures sensitive to sample and population size using of a positive real-valued parameter λ. First, however, I must discuss how Carnap dealt with languages with more than one observation predicate.

State descriptions can be defined for languages with more than one predicate. A state description in a language with k observational predicates $B_1,\ldots,B_k$ and n individuals, ($\mathscr{L}(B_1,\ldots,B_k, n)$) is one of the sentences of the form

$$(\pm B_1(a_1)\wedge\ldots\wedge\pm B_k(a_1))\wedge\ldots\wedge(\pm B_1(a_n)\wedge\ldots\wedge\pm B_k(a_n))$$

As before, these sentences serve as complete descriptions of states of affairs expressible in the language. Carnap developed a simpler means of expressing state descriptions using Q-predicates. A Q-predicate is a predicate of the form $Q_i(x) = \pm B_1(x)\wedge\ldots\wedge\pm B_k(x)$: it states for each property whether the individual possesses it or not. (Carnap denoted the number of Q-predicates in a language κ, which, in a language with n predicates is 2^n.) A state description in terms of Q-predicates is a sentence of the form $Q_i(a_1)\wedge\ldots\wedge Q_i(a_n)$.

Carnap showed that it was possible to express any open sentence in $\mathscr{L}$ using disjunctions of Q-predicates. For example, if we have a language with three observational predicates P_1, P_2, P_3, that is, $\kappa=2^3$, then $P_1(x)\wedge P_2(x)$ can be expressed by the disjunction of the Q-predicates $Q_1 = P_1(x)\wedge P_2(x)\wedge P_3(x)$ and $Q_2 = P_1(x)\wedge P_2(x)\wedge\neg P_3(x)$. Carnap denoted the minimum number of Q-predicates needed to express an open sentence as w, for 'width', and defined the relative width of a sentence as w/κ. Relative width expresses the degree of decomposability of a proposition, and Carnap employed it to avoid the paradoxes of the Principle of Indifference (indecomposable propositions have the smallest relative width).

The λ-continuum of inductive methods is that set of methods determined as follows: Given that r individuals in a sample of size k have A, the probability of the next observation being an A is

$$\cancel{X \to Y \to H_Y \boxminus H_X \wedge P_{\varepsilon_i}} \qquad \frac{r + \lambda(\frac{w}{\kappa})}{k+\lambda}$$

Thus the larger λ, the less sensitive the measure to relative frequencies. That is, the probability assigned to a statement about the relative frequency of an attribute in a set of observations is closer to the observed relative frequency the smaller the value of λ. To illustrate this, take the simple case of $\mathscr{L}(A,n)$, ($\kappa=2$ and $w=1$). This language is appropriate for describing a binomial trial. When $\lambda=\infty$, $c^{\infty}(A(a_{k+1})|e(k,r)) = 1/2$, no matter what the observed relative frequency, i.e., you never learn from experience. c^{∞} is, in fact, $c^{\dagger}$. When $\lambda=0$, $c^0(A(a_{k+1})|e(k,r)) = r/k$, which is the so-called straight rule (that is, the probability of the binomial mean of an attribute in a population is equal to the observed relative frequency of that attribute), and when $\lambda=2$, $c^2(A(a_{k+1})|e(k,r)) = (r+1)/(k+2)$ (c^2 is c^*.)

Carnap seems to have hoped to justify setting $\lambda=2$. However, I shall argue in the next section, he did not succeed in justifying this, or indeed any other, value for λ.

E. The Failure of the Analytic Method

We are now in a position to demonstrate that Carnap's attempt to provide an apriori basis for an inductive calculus fails. First, any measure c^λ is dependent on the language in which it is expressed since it depends on w and κ, the values of which depend on the number of observational predicates in our language, that is, on how we choose to classify observations. But the choice of a language, as Carnap pointed out, is not a matter for apriori considerations, but is determined by empirical and conventional considerations. Thus the choice of a measure is either empirical or conventional, not apriori.

Second, Carnap can give no apriori grounds for choosing one value of λ over another. For example, as mentioned above, Carnap rejected c^∞ (i.e., $c^\dagger$) because its values are unaffected by evidence that seems clearly relevant:

> ...this method ($\lambda = \infty$) clings always to that value of c... which is determined before any factual information is available. This value is never changed, no matter what factual information is obtained (as long as it does not concern the individual mentioned in the hypothesis). Thus accepting this method means refusing to give any regard to experience, to the results of observations, in making expectations or estimations. This is in gross contrast to what is generally regarded as sound inductive reasoning. (1952, p. 38)

Carnap rejected $\lambda=0$, the straight rule, advocated by Reichenbach, because it leads to counter-intuitive results. Taking again $\mathscr{L}(A,n)$, if our first observation reveals that $A(a_1)$, then by the straight rule, the probability that the next observed instance has A is 1. ~~If the~~ first observation is $\neg A(a_1)$, then the probability of the next observed instance being an A is 0. ~~In fact, if we are testing the universal statement that all a_is are A, then given the occurrence or non occurrence of $A(a_1)$, the probability of this universal law is 1 or 0. Since once a proposition is assigned probability 1 or 0, further observations cannot change its probability, this means that, on the basis of one observation, the straight rule confers certainty for a universal law~~. Carnap noted that this is the same as offering betting odds of 1:0 on the basis of one observation, which he took to be absurd: "Thus this method tells you that if you bet in the case described any amount on either of the two predictions mentioned, while the stake of your partner is 0, then this is a fair bet. I wonder whether there is any sane man, whether scientist, businessman, or gambler, who would actually regard a bet of this kind as fair." (1952, p. 43)

However, Keynes criticised the Rule of Succession, Carnap's preferred measure c^*, on similar grounds. He referred to "the absurdity of supposing that the odds are 2 to 1 in favour of a generalisation based on a single instance—a conclusion which this formula [the rule of succession] would seem to justify." (1921 [1973], p. 30, note 1) To the degree that Keynes and Reichenbach were sane men, it seems that Carnap is incorrect in his apriori restriction of probability measures.

Also, the λ-system is quite restricted in that, for example, it does not allow for probabilities of universal laws decreasing on positive instances of that law (i.e., counter-inductive measures). Carnap acknowledged this, but gave no apriori justification for classifying inductive methods according to the parameter λ: he relied instead on the intuitions of inductive logicians:

> ...we intend to be rather liberal in the admission of inductive methods to the projected λ-system. We shall not restrict it to those methods which appear plausible to a majority of thinkers but shall also admit methods which I and presumably many others would regard as unsatisfactory. On the other hand, we intend to exclude those methods which practically everybody would reject. (1952, p. 24)

Carnap said that "[n]o one would even consider" counter-inductive measures, claiming that "their use would be in conflict with the implicitly accepted basic principles of inductive reasoning":

> No attempt is made to include in the [λ-]system all conceivable inductive methods. This would be useless, because most of the methods which one could arbitrarily construct... would be entirely inadequate for the purpose of inductive application. No one would even consider them as possible methods of confirmation or of estimation because their use would be in conflict with the implicitly accepted basic principles of inductive reasoning (e.g., a kind that its value would be lower, the higher the rf [relative frequency] of [observations of a given attribute] M in the sample...). (1952, p. 8)

However, imagine the following scenario: we are observing some type of bird which is normally black, and we wish to the determine the probability that the next bird will be black. Carnap seems to have believed that it would be absurd to adopt a measure that would assign a lower probability to the event that the next bird will be black the more black birds we have observed. But, as Jon Dorling has pointed out (in conversation), if we take into account the well-known phenomenon of albinism in large animal populations, it might be that the more black birds we see which are not albinos, the higher the probability that the next bird will be an albino. Hence, it would in fact seem reasonable, in this case, to adopt an inductive method of the kind that Carnap rejected.

We may conclude, therefore, that Carnap was not able to justify apriori a unique inductive logic through the analysis of the meaning of the terms of the probability calculus.

In the first place, the inductive logic he constructed was dependent on a choice of language, and so was dependent on empirical considerations or conventions. Second, he could not justify preferring any one probability measure rather than another: Carnap instead seems to fall back on intuitions to adjudicate among the competing methods. And finally, his choice of inductive measures was too restricted. The first of these objections, is however, the strongest, and demonstrates clearly why there can be no inductive logic justified in an analytic manner: the results of analysis are dependent on the strength of the analysed language. However, the choice of a stronger or weaker language depends on empirical or conventional considerations, and hence there will be no one inductive logic obtained by taking the probability relation as purely analytic: empirical considerations always enter.

However, the quest for apriori foundations for inductive methodologies did not end with Carnap. Other writers dropped Carnap's symmetry conditions, and tried to establish weaker probabilistic methodologies on more fundamental principles. I refer to the Bayesian approach, and I shall summarise these attempts in the next chapter, and argue that they face the same difficulties.

CHAPTER 3: The Dutch Book Argument and Utility Theory as Apriori

The programme of establishing the probability calculus as analytic or intuitively obvious continued on lines different than those of Carnap and Keynes. I refer to the attempt to establish the probability calculus as providing obviously reasonable constraints on behaviour. This divides into two main strands: Dutch Book arguments, in which the probability calculus is claimed to provide consistency in betting situations; and utility theory, through which the probability calculus is constructed from restrictions on preferences. The constraints on betting behaviour and on preferences are taken as obvious in the sense that to adopt a course of action which would lead to a sure loss of money is clearly irrational. It is argued that only if the probability calculus is adopted as a guide to behaviour under uncertainty can this be avoided. I shall argue, however, in spite of the strong appeal of these arguments, that they are wrong and conclude that they cannot be interpreted as apriori approach. I conclude this chapter (and this Part of my thesis) with an argument that there can be not adequate apriori approach to methodology.

A. The Dutch Book Argument

Naturally, to discuss the Dutch Book argument I must first present it, which I do in the next section. In the following section I argue that the argument is unsound, or that it is invalid. I shall also argue that no plausible, valid formulation of the argument can be given.

1. The Argument Presented

I follow Skyrms 1986 and Howson and Urbach 1993 in my presentation of the argument. The Dutch Book argument employs the notion of betting, and concludes that fair bets must obey the probability calculus (betting behaviour is also taken to measure partial belief). A bet *on* a proposition A is an situation in which the bettor gets something of value, say, a certain sum of money a, if A comes to be accepted as true, and forfeits something of value, b otherwise. (A bet *against* A is an arrangement by which if A is accepted as false, the bettor against gets b, and loses a otherwise.) This can be represented by a "payoff table":

A	Payoff
T	$+a$
F	$-b$

(The above table represents a bet on A.) The *stake* S is defined as $a+b$ (i.e., the total amount of money at stake), and the odds b/a (so, for example, 4:1 odds). To simplify matters, I use the standard device of mapping the odds scale onto the interval [0,1] by setting $p = (b/a)/(1+b/a)$, which is $b/(a+b)$. p represents the normalised odds: this device is useful for showing how probabilities can be seen as constraints on bets. It also makes clear that p is a price rate for the purchase of a bet, that is, p is the proportion of the possible winnings paid in advance to take part in the betting. Employing these notions, the above payoff table can be rewritten as

A	Payoff
T	S(1-p)
F	-Sp

A value for p which is given in the belief that it confers no advantage to either side of the bet is called the *fair betting quotient.* In other words, if the purchaser of a bet believes that p is a fair betting quotient then he should be indifferent to betting on or against the proposition. Consider a bet on A in which $p<0$. As is clear, in a bet on A, regardless of the truth or falsity of A, the payoff on is positive, and the payoff against is negative. So the bettor on will always make a gain, and the bettor against will always make a loss. Hence, a bet in which $p<0$ is not fair: anyone who regarded this as a fair betting quotient would be indifferent between a sure loss and a sure gain. Such a bet, however, is clearly not fair. This concludes the argument that

(i) For all A, $p(A) \geq 0$.

The force of the argument is that to be indifferent between a sure loss and a sure gain is in some sense irrational or at least wrongheaded, and that this irrationality can be avoided by giving non-negative betting quotients.

A similar argument can be made concerning a bet on a tautology T. Since T is necessarily true, we need only consider the payoff S(1-p). If $p>1$, then the bettor on will necessarily make a net loss and a bettor against a net gain. If $p<1$ the bettor against will likewise always make a net loss, the bettor on a net gain. Thus, $p=1$ represent the only fair odds on T. Hence,

(ii) For all tautologies T, $p(T) = 1$.

The final argument concerns three separate bets on the propositions A, B, and $A \vee B$, where A and B are mutually exclusive, that is, $A \vdash \neg B$. If the betting quotient for A is p_1 and for B is p_2 then the payoff table for the combined bet on A and on B is:

A	B	Net Payoff
T	F	$S(1-p_1) - Sp_2$
F	T	$-Sp_1 + S(1-p_2)$
F	F	$-Sp_1 - Sp_2$

But this is just a bet on A∨B with a betting quotient p_1+p_2, as can be seen by rewriting the above payoff table setting $p_1+p_2 = r$

A	B	Net Payoff
T	F	$S(1-r)$
F	T	$S(1-r)$
F	F	$-Sr$

which is the same as

A∨B	Payoff
T	$S(1-r)$
F	$-Sr$

Hence the two separate bets on A and on B are the same as the single bet on A∨B with a betting quotient p_1+p_2. Suppose, however, that someone offers betting quotients p_1, p_2 and r on A, B and on A∨B such that $p_1+p_2 \neq r$. Then, by definition, that person would be indifferent between betting on A and on B and against A∨B at those rates. The payoff table for this set of bets is

A	B	Net Payoff
T	F	$S(1-(p_1+p_2)) - S(1-r)$
F	T	$S(1-(p_1+p_2)) - S(1-r)$
F	F	$-S(p_1+p_2) + Sr$

which is the same as

A∨B	Net Payoff

T	$S(r-(p_1+p_2))$
F	$S(r-(p_1+p_2))$

If $r>p_1+p_2$, this set of bets would lead to a sure gain, and if $r<p_1+p_2$, to a sure loss. It is clear from the above payoff table that the bets can only be fair when $r=q$. Hence, if A and B are mutually exclusive, their betting quotients must add to be fair:

(iii) $p(A \vee B) = p(A) + p(B)$, if A and B are mutually exclusive

This completes the Dutch Book arguments for the Kolmogorov axioms for unconditional probability, and finite additivity.

There is a Dutch Book argument for the fourth axiom. However, my arguments later will not deal with this axiom, and so I shall not discuss conditional probabilities in the rest of this chapter. Also, all that has been argued for so far is that adherence to the axioms of probability is a *necessary* condition for fairness (that is, if a set of betting quotients do not adhere to the probability calculus, then they are not fair, or conversely, if a set of betting quotients is fair, then they adhere to the probability calculus). There is an argument that adhering to be shown that it constitutes a *sufficient* conditions for fairness (that is, if the betting quotients adhere to the probability calculus, they are fair). I shall not present this, since my later arguments will not turn on this aspect of the Dutch Book argument. (The proof can be found in Howson and Urbach 1993, pp. 84-85.)

The modern formulation of the Dutch Book argument stems from a series of papers by Shimony (1955), Lehman (1955) and Kemeny (1955). At least one of these authors, Shimony, aimed to provide apriori grounds for a probabilistic methodology. He explicitly mentioned the difficulties with Keynes's intuitionistic justification: "Both Keynes and Koopman justify their axiomatizations by a claim of self-evidence. However, if it is meaningful to speak of degrees of self-evidence, many of Keynes's axioms and several of Koopman's... are less self-evident than is desirable. Consequently, a more adequate justification of the axioms... is needed." (1955, p. 4) Shimony claimed that the Dutch Book argument established the axioms of the probability calculus as analytic: "The Principle of Coherence [i.e., the relation between fair betting quotients established by the Dutch Book argument] seems to be analytic; i.e., it is true in virtue of relations which hold between the concepts of confirmation and the concepts of coherence, rather than in virtue of any contingent facts." (1955, p. 10-11) (However, Shimony did not seem to believe that the argument was analytic in the Carnapian sense, since he felt that it was the type of argument which "assert[s] such fundamental relations among the concepts involved in them that in defending them, one can only appeal to an intuitive grasp of these concepts. (1955, p. 11-12))

Bruno de Finetti, perhaps the first to point out how a Dutch Book argument could be constructed, also seemed to believe that the probability calculus was in some sense analytic. de Finetti dismissed empirical considerations as irrelevant when considering whether or not the probability calculus should be a guide to uncertainty:

> Experiments of this kind, which are made in order to check the extent to which actual behaviour conforms to the norms derived from the theory of probability, are often considered as 'proving' or 'disproving' the validity of probability theory (or the related theory of decision-making under conditions of uncertainty). This would be so if such theories were to be regarded as empirical-psychological theories of actual behaviour, but, in fact, it is completely at odds with what we are considering here: a *normative* theory for *coherent* behaviour. (De Finetti 1974, p. 192)

Instead, De Finetti seemed to believe that the foundations of probabilistic reasoning are like those of mathematics, in that they are immune to revision in the light of empirical evidence:

> This kind of empirical evidence is also of interest from our standpoint, but in the same way as a mathematician might find the mistakes of laymen, students, or even other mathematicians interesting. He does not modify mathematics by incorporating these 'mistakes', as though, simply because someone has enunciated them, they 'should' be included by virtue of their being part of some psychological truth, or of the indiscriminate collection of mathematical statements made in the course of history. (ibid)

In contrast to this, I shall argue in the next section that not following the probability calculus is not the same as making a mistake in mathematics. Indeed, I shall argue that the Dutch Book argument is either unsound or invalid, depending of the formulation. Indeed, I shall give a general argument that no valid and sound formulation of the argument can be given.

2. The Dutch Book Argument is invalid

In this section I shall offer two different interpretations of the Dutch Book argument which I call, for reasons which will become obvious, behaviourist and counterfactual. I argue that if we adopt the first interpretation, the Dutch Book argument is unsound, and that if we adopt the second, it is invalid. I shall then argue that it is not possible to formulate the Dutch Book argument in any way in which it is neither unsound nor invalid. This will conclude my discussion of these types of arguments, and in the next section I shall argue consider utility theory.

It has been argued (by Schick in his 1986 "Dutch Bookies and Money Pumps") that the Dutch Book argument for the third axiom requires the assumption that the value of money is additive. That this assumption is indeed required can be seen by considering the following example. Suppose that I have a set aside a portion of my (limited) income, and am considering buying a bet on A, which I believe quite strongly, but not with certainty, to be true. If I were to only buy the bet on A, then I would pay a high price, that is, I would give a high betting quotient. But now suppose that I have the opportunity to by a bet on a proposition B, which is mutually exclusive with A. It seems natural that I would, in this circumstance, be less willing to pay the same amount for A if bought as part of a package of bets than I would be willing to spend on A if bought it alone. In other words, the prices I pay for bets bought jointly might turn out to be subadditive. I suggest that this is how most people would feel: they wish to pay less for items bought in bulk.

Schick labels the assumption that people are willing to offer the same betting quotients singly that they would offer jointly 'value additivity.' Another, more traditional, way of putting the objection is that the value of money is not linear. It has been pointed out on many occasions that the Dutch Book argument assumes the value of money to be linear. This criticism seems to me correct: the fair betting quotient p is a ratio derived from a particular betting situation in which the bettor is willing to risk forfeiting some amount of money b to get an amount a if the proposition being betted on is true. However, it is a strong assumption that the proportion $b/b+a$ will remain the same for all possible values of a and b. Most people would offer much more conservative rates the larger the amount of money involved. It seems, then, that the Dutch Book argument rests on a demonstrably false assumption of the linear value of money.

However, this interpretation of the Dutch Book argument might seem to be overly behaviourist: it requires that people actually be willing to place money for a bet, instead of merely naming some quantity which they think is a fair price. Howson and Urbach have suggested that a better interpretation of the Dutch Book argument would be a counterfactual one: the betting rate p is the rate that one *would* be prepared to offer if one *were* to bet. At first glance, this seems superior to the behaviourist approach:

> Attempts to measure the values of options in terms of utilities are traditionally the way people have sought to forge a link between belief and action, and much contemporary Bayesian literature takes this as its starting point. We do not want to deny that beliefs have behavioural consequences in appropriate conditions, they clearly do, but stating what those conditions are with any precision is a task fraught with difficulty, if not impossible... [T]he conclusion we want to derive, that beliefs infringing a certain condition are inconsistent, can be drawn merely by looking at the consequences of what *would* happen if anyone *were* to bet in the manner and in the conditions specified. (Howson and Urbach 1993, p. 77)

Problems about the non-linear value of money are avoided, as are other problems such as bets on propositions which cannot be decided in a finite amount of time, or unwilling subjects who object to betting, since no money actually changes hands. The Dutch Book argument is, under this interpretation, conceived of as a thought experiment, establishing rules of consistency of ideal behaviour.

Unfortunately, if we employ the standard Lewis-Stalnaker semantics, this interpretation of the Dutch Book argument makes it invalid. (This argument was suggested to me by a similar argument by Anand 1993, which I shall detail in section B.2). This can be seen by formulating the Dutch Book argument for the third axiom as follows.

(1) If you were to bet on A you would regard p as a fair betting quotient.

(2) If you were to bet on B you would regard q as a fair betting quotient.

Therefore

(3) If you were to bet on A and on B, you would regard p and q as fair betting quotients.

Using the usual symbol for the counterfactual conditional we may rewrite the argument

(1) $A \mathrel{\Box\!\!\rightarrow} p$

(2) $B \mathrel{\Box\!\!\rightarrow} q$

Therefore

(3) $A \wedge B \mathrel{\Box\!\!\rightarrow} p \wedge q$

This is an invalid pattern of reasoning in Lewis-Stalnaker semantics: it is an instance of the so-called counterfactual fallacy of strengthening the antecedent:

$A \mathrel{\Box\!\!\rightarrow} p$

$\therefore A \wedge B \mathrel{\Box\!\!\rightarrow} p$

(A discussion of strengthening the antecedent of counterfactual conditionals can be found in Lewis 1973 [1986], p. 17.) Hence, the most plausible formulation of the Dutch Book argument is invalid.

It might be argued that the argument could be formulated in another logic: counterfactual conditionals with Lewis-Stalnaker semantics just might be not be the correct way to formulate the argument. I shall now argue that any logic which plausibly formulates the Dutch Book argument will also show the argument to be invalid. The behaviourist formulation of the Dutch Book argument is unsound because it rests on assumptions about dispositions to economic behaviour of a certain sort. These dispositions are often not realised, due to various economic factors: perhaps the bettor worries about spending money, or is morally opposed to betting, or some similar accompanying disposition. The counterfactual approach was introduced to dispense with these other, interfering, factors. However, Lewis-Stalnaker semantics make the argument invalid. And, indeed, any logic which faithfully represents dispositions to behaviour will also make the argument invalid:

dispositions are in general not serially or jointly realisable, as the following examples shows. Suppose that I actually live in Prague, and it is summer. Further suppose that I am feeling suicidal (unlikely, granted, since I live in Prague: feel free to substitute another city with the necessary conditions) and that I am disposed to have a beer, because it is hot. But it is surely possible that if I drink the beer first, I shall no longer feel suicidal, because it is so good, or perhaps because I get too drunk to commit suicide. Conversely, suicide precludes beer drinking. Clearly my beer drinking disposition blocks my disposition to commit suicide. This seems to be the case in general with dispositions, and in particular with dispositions to economic behaviour, as the non-linear utility of money shows.

I conclude that the Dutch Book argument, in its most plausible formulation, is ~~irredeemably~~ invalid. Of course, someone might choose the less plausible behaviourist formulation. But in this case, the argument is surely not apriori. This is not to say that following the probability calculus as a guide to behaviour uncertainty is under no circumstances justified: it is to say that it is not justified apriori by the Dutch Book argument. However, the Dutch Book argument is not the only argument available for probabilistic methodology. In the next section I consider the simultaneous construction of utilities and probabilities from restrictions on preferences. I shall also argue that these restrictions are not apriori justified.

B. Probabilities from Preferences

In this section I shall discuss the derivation of probabilities in utility theory. I shall not, however, introduce all the axioms of the possible systems, nor shall I examine any particular theory in depth. I shall merely note that all these theories require that our preferences be transitive (the technical reason for this is to establish a partial ordering of probabilities and utilities: a linear order is introduced by independence assumptions). I shall argue in this section that the requirement that our preferences be transitively ordered is not apriori justified. I shall thus conclude that utility theoretic constructions of probability are not justified apriori.

Transitivity is a natural restriction on preferences.[2] It is seemingly obvious that we should not have cyclical preferences; that is, that the following should hold:

$$B \succ A \wedge C \succ B \Rightarrow C \succ A$$

L.J. Savage, who famously developed the utility theoretic approach to subjective

[2]The usual notion for "A is preferred to B" is "$A \succ B$" ("strong preference"). Indifference between A and B is denoted by "$\sim$" while weak preference (preferred or indifferent) is denoted by "$\preceq$."

probability in his *The Foundations of Statistics*, advocated a distinction between normative and empirical accounts of probability. For Savage, transitivity of preferences (his postulate P1) was a normative, and not an empirical criterion of behaviour:

> Two very different sorts of interpretations can be made of P1 and the other postulates to be adduced later. First, P1 can be regarded as a prediction about the behavior of people, or animals, in decision situations. Second, it can be regarded as a logic-like criterion of consistency in decision situations. For us, the second interpretation is the only one of direct relevance, but it may be fruitful to discuss both, calling the first *empirical* and the second *normative*. (Savage 1972, p. 19)

Savage held that transitivity was an intuitively rational constraint on preferences. He argued that intransitive preferences are like contradictions in logic, as can be determined by introspection:

> ...when it is explicitly brought to my attention that I have shown a preference for **f** as compared with **g**, for **g** as compared with **h**, and for **h** as compared with **f**, I feel uncomfortable in much the same way that I do when it is brought to my attention that some of my beliefs are logically contradictory. Whenever I examine such a triple of preferences on my own part, I find that it is not at all difficult to reverse one of them. In fact, I find on contemplating the three alleged preferences side by side that at least one among them is not a preference at all, at any rate not any more. (ibid, p. 21)

The requirement of transitivity of preferences is not only obvious, but useful: transitivity of preference is an axiom in all utility theoretic constructions of (standard, real-valued quantitative) probability. However, an argument is needed to establish this restriction as apriori, for, as I have argued in the first chapter, appeals to intuitive obviousness does not establish claims apriori.

A seemingly powerful argument is available: the Money Pump argument. Like the Dutch Book argument, it has both a behaviourist and a counterfactual interpretation. I shall discuss these in turn, after presenting the argument.

1. The Money Pump argument

It seems reasonable to assume that to have a preference for some object X over another object Y means that, if you have Y, you will be willing to pay some amount of money, no matter how small (which I shall denote by ε) in order to exchange Y for X. But suppose you have intransitive preferences over three objects A, B and C, i.e.,

$$C \succ B,\ B \succ A,\ A \succ C.$$

The Money Pump argument goes thus: if you have A then you will be willing to pay exchange it for B. And, when you get B then you will be willing to pay to exchange it for C. But when you have C, you will be willing to pay to get A, which you began with. So, if you have intransitive preferences, you will spend money to get what you previously had. Even worse, there is no reason, unless you change your preferences, for this cycle to stop. Hence, arbitrarily large amounts of money may be had from you—you may be pumped of all your resources, hence the name "The Money Pump Argument." Being prepared to be pumped dry clearly seems to be irrational behaviour.

2. The Money Pump argument is invalid

However, the argument just given is not complete: the structure of the currency used in the exchanges is not defined. Our currency may be continuous (infinitely divisible) or discrete (this corresponds to a very large and a small initial fortune, respectively). We can denote the amount of money spent for the *i*th exchange as ε_i. If the currency is continuous, we could choose the ε_is such that

$$\lim_{i \to \infty} \sum_i \varepsilon_i = \delta$$

for some extremely small δ. It might be claimed that this is in fact a pump. However, it is clearly no sufficient motivation for having transitive preferences. If money were like dust, then it does not seem irrational to throw some away. Or, millionaires who light cigars with one hundred dollar bills surely act in an odious fashion: but it does not seem that they act irrationally.

It seems more promising to consider a discrete currency: however, the argument is then unsound. As formulated above, it follows from having a preference between two objects there will always be some amount of money which we would be willing to exchange the less preferred object for the more preferred. This, however, does not seem to be the case. It seems possible to have a preference, but not be willing to spend a full unit of currency to realise it: I might prefer one object over the other, and be willing to spend one-hundredth of a cent. However, since this currency is not available, I would not be willing to exchange. Hence, my preferences, although intransitive, would not be elicited, and so I would not be pumped. The Money Pump argument shares the same difficulties as the Dutch Book argument: in its behaviourist formulation it is valid, but unsound. The counterfactual formulation is invalid, as I shall now show.

Using counterfactuals, the Money Pump argument can be stated as follows:

If I had C, then I would be prepared to pay ε to exchange it for B

If I had B, then I would be prepared to pay ε to exchange it for A

If I had A, then I would be prepared to pay ε to exchange it for C

Therefore

If I had C, then I would be prepared to pay $\varepsilon_1+\varepsilon_2+\varepsilon_3\ldots$ to exchange it for C. To formalise the argument, I use the predicates H for "to have x" and P for "to pay ε_i". Using counterfactuals, the notion of preference can be formulated as follows:

$$X \succ Y \Rightarrow H_Y \Box\!\!\rightarrow H_X \wedge P_{\varepsilon_i}$$

(In fact, this formulation is incorrect since it does not capture Y's being exchanged for X. This can be solved by indexing the predicates by time. However, I leave this out for clarity's and brevity's sake). The money pump argument then becomes

$$H_C \Box\!\!\rightarrow H_B \wedge P_{\varepsilon_i}$$
$$H_B \Box\!\!\rightarrow H_A \wedge P_{\varepsilon_i}$$
$$H_A \Box\!\!\rightarrow H_C \wedge P_{\varepsilon_i}$$

(The ε_i are not necessarily the same.) To deduce a pump, however, would require that counterfactual conditionals be transitive, which they are not under Lewis-Stalnaker semantics. (This was first pointed out by Anand 1993.) It might be argued, again, that counterfactual conditionals do not appropriately represent dispositions. But, as in the case of the Dutch Book argument, any logic which is adequate for dispositions will not have transitivity as a valid rule of inference, since, as the example of beer drinking and suicide shows, dispositions are not, in general, serially realisable.

I conclude that, like the Dutch Book argument, the Money Pump argument, if valid, is unsound, and if sound, invalid. Of course, there are many arguments for probabilistic methodologies which are claimed to be apriori: it is simply not possible to canvas all such argument, much less all possible arguments. Nonetheless, I have attempted to show that the two most popular arguments for probabilistic methodology are not apriori. In the next section I argue that there is reason to conclude that all attempts to establish a particular methodology as apriori are bound to fail.

C. The Failure of the Apriori Approach to Methodology

In the first part of this thesis I have considered three different types of approaches to apriori methodology: intuitionism, logical analysis, and consistency arguments. I have argued that they have all failed. However, I have not shown that all apriori approaches must fail. And it is surely not possible to do so. I shall argue in the next chapter that the mere fact that all attempts to establish methodology apriori have failed up to this time does not imply that all attempts afterwards shall likewise fail. But I shall now argue that all apriori approaches suffer from the same defect as intuitionism: that it is always possible to reasonably disagree with them.

Robert Nozick (in the introduction to his 1981) considers what seems to be the strongest argument possible: one so powerful, that if someone did not accept it, they would die. But, as he notes, someone might choose not to accept the argument, and die nonetheless. It might seem silly to reject such an argument, but that does not establish the argument as apriori correct. Indeed, it seems to always be possible to reject any argument, and this is as it should be. An argument which may seem obvious may later turn out not to be (as did, for example, Moore's claim that the good was non-natural and undefinable, or Carnap's acceptance of symmetry principles). Indeed, it seems in keeping with being a limited biological organism: we can never know what the 'absolute truth', if there is such a thing, is. In this part, I have rejected two different apriori approaches: intuitionism and the method of logical analysis. In the field of methodology, these two approaches seem to exhaust purely apriori methodologies. I have also argued that they should be rejected. This of course does not prove that no apriori approach will be discovered, although it does seem to show that an aprioristic approach to methodology is, generally, a bad idea. And, further, my arguments shift the burden of proof onto the advocates of the apriori approach. As I shall show in the next chapters, the advocates of non-empirical methodology equate 'normative' with 'apriori.' Before this may be done, however, an adequate apriori approach needs to be developed. Until the time that one is, mere assertions of the apriori character of normativity will have no force.

As I said above, I doubt such an adequate account can be developed. Broadly stated, the problem with the purely aprioristic approach for methodology is one of epistemic access: there just does not seem to be a way to guarantee that we can know an apriori correct methodology if presented with one. Whatever the merits of a purely apriori approach in mathematics, it seems out of place in methodology. Thus, it seems pragmatically better to adopt a more cautious approach to methodology: I shall outline this approach in the final chapter. In the next chapter I shall consider a purely empirical approach to methodology. I shall conclude that, like the purely apriori approach, it is doomed to fail, for much the same reasons.

PART II: METHODOLOGY AS EMPIRICAL

I have so far discussed the claim that methodology is apriori, and not empirical, and I have argued that this point of view is inadequate. One widely proposed alternative is that methodology should be, at least to some degree, an empirical study: this position, as noted in the introduction, is usually called naturalism. The degree to which methodology is an empirical, as opposed to an apriori undertaking is a matter of debate between those who adopt an empirical point of view. In Chapter 4, I will discuss arguments in favour of a purely empirical approach, and conclude that these arguments are not successful. I shall argue that there are powerful arguments against such a position. These arguments motivate the search for a partly empirical, partly normative approach to methodology. In Chapter 5, I interpret Quine as offering such an approach, but conclude that he is not successful in arguing for it.

CHAPTER 4: The Purely Empirical Approach

One alternative to the position that methodology should be non-empirical is commonly referred to as naturalism. I shall take the term "naturalism" as a convenient abbreviation of "empirical approach to epistemology." There seem to be many types of naturalism (since, indeed, there are many approaches to empiricist epistemology.) In this chapter I consider 'pure' naturalism: the view that methodological statements are not apriori, but aposteriori. The two most famous philosophers associated with this position are W.V.O. Quine, who, it is often claimed, advocated this position in his 1969 "Epistemology Naturalized", and Ronald Giere. I intend to argue in the next chapter that Quine does not endorse the position commonly attributed him, but for the purposes of discussing the purely empirical approach I will refer to Quine's views as they are standardly interpreted.

A. The Argument From the Failure of the Apriori Programme

Very few arguments seem to have been put forward for the purely empirical approach. In this section I discuss the argument for a purely descriptive account of the methods of science from the failure of the apriori programme. A similar argument has also been advance by Giere, and I shall also discuss it.

The argument from failure begins with the claim that repeated attempts over many centuries to formulate a satisfactory apriori methodology have been unsuccessful. Science, however, it is claimed, has during this time, been successful. Therefore, it is argued, philosophers should adopt the successful empirical approach in place of the unsuccessful apriori programme. There are two main versions of this argument, one attributed to Quine, the other to Giere. I shall discuss these in turn.

Quine, as standardly interpreted, offers only the argument from failure for his position in "Epistemology Naturalized," which is said to run as follows: Epistemology has progressed (or rather, not progressed) by successive failures, and this shows that non-empirical approaches cannot succeed. The conclusion is that a purely descriptive approach is the only alternative for epistemologists. Quine discusses Descartes, Hume, Russell and Carnap, and concludes that their investigations demonstrate the futility of attempting to found epistemology on apriori grounds. On the other hand, Quine argues that is possible to explain how people form theories about the world using behaviourist psychology. He concludes that epistemologists should concentrate their efforts on the psychological study of the causal mechanisms that lead from sensory stimulation to belief, rather than on supposedly aprioristic accounts of theory evaluation. This substitution of psychology for apriori epistemology as the tool for understanding methodology is called the naturalisation of epistemology.

This is widely acknowledged as Quine's argument, e.g., by Hilary Kornblith, who in his introduction to the collection of essays *Naturalized Epistemology* gives exactly this reading of "Epistemology Naturalized":

> Quine's argument... is this: the history of epistemology is largely the history of the foundationalist program... The history of epistemology shows that the foundationalist program has faced one failure after another. The lesson to be learned from these failures, according to Quine, is not just that foundationalists had mistakenly answered... [the question "How ought we to arrive at our beliefs?"] in claiming that the appropriate way to arrive at one's beliefs is to begin with beliefs about which one cannot be wrong and build upon that foundation. Rather, according to Quine, foundationalists were asking the wrong questions. Once we see the sterility of the foundationalist program, we see that the only genuine questions there are to ask about the relation between theory and evidence and about the acquisition of belief are psychological questions. (Kornblith 1985, p. 3-4)

Putnam likewise concurs with the standard interpretation of Quine's argument:

> "Justification" has failed. (Quine considers the notion only in its strong "Cartesian" setting, which is one of the things that makes this paper puzzling.) Hume taught us we *can't* justify our knowledge claims (in a foundational way). Conceptual reduction has also failed (Quine reviews the failure of phenomenalism as represented by Carnap's attempt in the *Logische Aufbau*). So, Quine urges, let us give up epistemology and settle for psychology. (Putnam 1982, p. 298)

Jaegwon Kim also agrees with the standard interpretation: "Quine's principal argument in this paper against traditional epistemology is based on the claim that the Cartesian foundationalist program has failed—that the Cartesian "quest for certainty" is "a lost cause"." (1988, p. 385) Thus, Quine is interpreted as tracing the history of a particular branch of epistemology, and concluding from its failure that epistemology as aspired to by philosophers must be abandoned, and psychology put in its place.

The argument for empirical methods in place of apriori theorising from the alleged failure of normative methodology is also offered by Ronald Giere. Giere sees himself as arguing in the same manner as Quine did in "Epistemology Naturalized" (hence, presumably, the title of Giere's 1985 "Philosophy of Science Naturalized"), and stresses the failure of philosophers of science to explicate and define "rationality." Whereas Quine evidently assumes that the search for a justification of science ended with Hume, Giere extends the discussion to philosophers of science of this century, in particular, Carnap, Reichenbach, Popper, Lakatos, Laudan, and members of the subjective Bayesian school of methodology. Giere argues that non-empirical (normative) methodologies fail on one of three counts: they do not live up to the apriori claims made for them (Carnap, Reichenbach, Bayesianism), or they are merely descriptive, hence naturalistic (Lakatos, Laudan), or fail to describe properly scientific practice (Popper). (The last criticism is curious, since Popper has avowedly disowned descriptive approaches, as will be shown in Part III, Chapter 6.) Be that as it may, Giere argues that we should regard philosophy of science as descriptive and aposteriori.

However, Giere's reasons for adopting naturalism are not entirely clear. He claims that while the failure of current non-empirical approaches to methodology provides a reason or "motivation" for investigating a naturalistic approach, this failure does not prove that such an empirical investigation will succeed: "This [failure] hardly proves that there is no way to achieve these ends [i.e., an apriori justification of science]. It does, however, provide some motivation for seeking to *understand* how a naturalized philosophy of science might fruitfully be pursued." (1985, p. 339)

Giere's argument can be read in two ways (as, indeed, may Quine's). The first is as an apriori argument for an aposteriori approach to methodology. Thus Giere gives apriori reasons for rejecting the methodologies proposed by other philosophers, and argues that these are not applicable to his methodology. Or, Giere's argument can be read as an empirical argument for an empirical approach, that is, as an inductive argument that non-empirical philosophy of science is discredited by its failures, and that empirical approaches are much more promising, in light of their success. The second interpretation is supported by Giere's other writings, for example, in his *Explaining Science*:

> For any given theory and any given set of data, there is said to be a "rationally correct" conclusion about the extent to which those data "rationally support" the theory. The philosopher's task has been seen as one of making explicit the principles that scientists are deemed to employ intuitively in evaluating theories and then to show that these principles are indeed rationally correct.
>
> The major philosophical problem with this approach has always been to demonstrate that a particular principle captures a relationship that is uniquely rational. In fact, addressing that problem is what most of the literature is all about. The inconclusiveness of the literature, after many years of effort, is generally taken to indicate the difficulty of the problem. *But it may also be taken as a basis for suspecting that there is something fundamentally mistaken about the whole enterprise.* (Giere 1988, p. 3, my italics)

Later in the same book, Giere repeats this argument, claiming that the failure of apriori justifications of science gives support to the claim that such attempts should be abandoned:

> The programme of trying to justify science without appeal to any even minimally scientific premises has been going on without conspicuous success for 300 years. One begins to suspect the lack of success is due to the impossibility of the task. Perhaps there is simply no place totally outside science from which to justify science. At the very least one might conclude that the task is not going to be accomplished anytime soon by ordinary mortals. That would seem to be sufficient grounds for an ordinary mortal to try something else. (Giere 1988, pp. 11-12)

Giere puts the same point in a slightly different way by arguing that empirical success will serve as the justification of a purely empirical methodology. He takes as an example of an apriori approach being rejected on aposteriori grounds the philosophical arguments of the scholastics concerning physics: he claims that the scholastic approach of apriori reasoning in physics was rejected not because their arguments were found to be flawed, but because modern physics was empirically successful. Presumably Giere thinks that the empirical approach to the philosophy of science will supersede the apriori approach in the same way:

> It [naturalism] is also part of a long tradition of using science in the attempt to understand science itself. Like all such general viewpoints, naturalism and the scientific study of science have been "refuted" by philosophers dozens of times. What makes naturalism attractive now, however, is not that such philosophical arguments have themselves been "refuted" but that the cognitive sciences have become increasingly successful empirically. This is the historical pattern. Proponents of the new physics of the seventeenth century won out not because they explicitly refuted the arguments of the scholastics, but because the empirical success of their science rendered the scholastics' arguments irrelevant. (Giere 1988, p. 8-9)

However, there is also textual support for the first interpretation of Giere's argument from failure. In Giere's 1989 "Scientific Rationality as Instrumental Rationality," he repeats the argument for the empirical approach from the failure of the apriori programme, and attributes this failure to a mistaken philosophical view of justification (what he calls "Aristotelian"). According to Giere, the apriori approach is committed to "an Aristotelian conception of *justification.*" According to this conception, justification "must proceed in a strictly linear fashion from principles that, on pain of regress or circularity, are self-evident or... 'self-justifying'." Giere then repeats his argument from failure:

> Apart from commitment to an Aristotelian picture of justification, what reason is there to think that there are any such principles? Three hundred years of epistemological enquiry by some of the finest minds in the Western intellectual tradition have failed to reveal any set of principles that could claim the allegiance of a majority even of philosophers. (Giere 1989, p. 378)

Giere acknowledges that the failure to produce an 'Aristotelian' justification so far does not demonstrate that no such justification is available: "Of course, by Aristotelian standards this failure does not *prove* that no such principles exist. Nevertheless, the failures of traditional epistemology up to now surely invite the suggestion that the project was misconceived from the start." (ibid) However, Giere approvingly mentions what he takes to be Richard Rorty's claim that "the original sin was to think that an Aristotelian sort of justification is even possible, let alone necessary." Giere does not pursue this line of reasoning, but the implication seems to be that there is an alternative to an "Aristotelian conception of justification." If my reading of Giere is correct, it shows that he holds, at least some of the time, that we can argue from an apriori standpoint for abandoning apriori approaches to methodology. The question of how Giere defends his position will be important in section B.3, where I discuss the claim that purely empirical approaches to methodology are self-refuting.

The argument from failure—that apriori epistemology has, in spite of considerable effort by well-qualified people, failed, and that we should therefore pursue epistemology as an empirical study—faces severe difficulties. The argument requires that the only alternative to apriori methodology be aposteriori methodology: for example, Giere claims that "[t]he *only viable philosophy of science* is a naturalized philosophy of science." (Giere 1985, p. 355, italics mine) But the conclusion does not follow from the premisses: it remains to be demonstrated that the only two possibly viable approaches to epistemology are apriori and aposteriori. I shall argue in the final two chapters that another alternative, namely conventionalism, is available (and superior). Another difficulty is in arguing for the success of science: it seems that one could argue that since all scientific theories are probably false, an empirical approach to methodology will also be probably false. I shall also argue in the following section that attempts to give a purely empirical account of scientific success result in circular arguments, or that they are self-refuting.

B. Arguments against Naturalism

Giere argues that we should not attempt to set out apriori principles of science, but should attempt to understand science in purely descriptive or empirical terms. Such an approach faces severe difficulties. The three most common criticisms are that the purely empirical approach can only proceed by changing or begging epistemological questions of justification, not by answering them; that it is self-refuting; and that it can only be argued for circularly. I shall argue that these criticisms pose a dilemma: either naturalism may be argued for on apriori grounds, in which case it is self-refuting, or it may be argued for on aposteriori grounds, in which case it is justified in a circular manner. (Of course, one could argue that there is no such thing as justification: but in this case, *a fortiori*, adopting a purely empirical approach would be no more justified than adopting an apriori approach.) Before I develop this dilemma, I will first address an argument against naturalism which

I shall argue is not successful: that adopting a descriptive approach to methodology is question begging.

1. Naturalism is changing the question

Several philosophers have complained that the adoption of a purely empirical approach is not a solution to problems of apriori epistemology, but a changing of the question. For example, Rorty accuses Quine of changing "the motive of inquiry." Quine suggests in "Epistemology Naturalized" that we look at the causal processes (such as sensory stimulation) which lead to beliefs. But for Rorty, an examination of how beliefs come to be held is not what philosophers are interested in. Philosophers are interested in whether beliefs are justified or not, and so Rorty accuses Quine of "changing the motive of inquiry." Rorty claims that epistemology is concerned with the justification of some theories over others, not in determining what causal mechanisms are involved in the development of these theories:

> If one were only interested in causal mechanisms, one would never have worried one's head about awareness. But the epistemologists who dreamed the dream Quine describes were not only interested in causal mechanisms. They were interested in making an invidious distinction between Galileo and the professors who refused to look through his telescope. (Rorty 1980, p. 225)

Rorty addresses Quine's claim that an empirical epistemology would be free of a "stubborn old enigma of epistemological priority." (Quine 1969, p. 84) The enigma is the following: observation might be said to begin with only two dimensions, the patern of light created by the object on the retina. However, we unconsciously translate this stimulation, and see things in three dimensions. Yet both the irradiation of the retinas, and the mental image which results seem to serve equally well as an explication of the term observation, hence the "enigma" or "dilemma" which Quine refers to: "Our retinas are irradiated in two dimensions, yet we see things as three-dimensional without conscious interference. Which is to count as observation—the unconscious two-dimensional reception or the conscious three-dimensional apprehension?" (Quine 1969, p. 84) Quine suggests that the irradiation of the retina is observation: "What to count as observation now can be settled in terms of the stimulation of sensory receptors, let consciousness [i.e., the mental image produced by the stimulation] fall where it may" (Quine 1969, p. 84). Rorty claims that by advocating such an approach, Quine advocates an abandoning, not naturalising, epistemology:

> If there are indeed no experimental criteria for where the real data come, then Quine's suggestion that we give up the notion of "sense data" and speak causally of nerve endings and epistemologically of observation sentences does not resolve a dilemma which has plagued epistemology. Rather it lets epistemology wither away. For if we have psychophysiology and history of science to note the

> occasions on which observation sentences are invoked or dodged in constructing and dismantling theories, then epistemology has nothing to do. (1980, p. 225)

Kim also charges Quine with changing the question. For Kim, empirical epistemology is an oxymoron: epistemologists are by definition concerned with justification of belief, empirical scientists are concerned only with belief formation. According to Kim the one studies justification, the other "causal-nomological" relations:

> To be sure, they both investigate "how evidence relates to theory." But putting the matter this way can be misleading, and has perhaps misled Quine: the two disciplines do not investigate the same relation... normative epistemology is concerned with the evidential relation properly so-called—that is, the relation of justification—and Quine's naturalized epistemology is meant to study the causal-nomological relation. (Kim 1988, p. 391)

Thus, according to Kim, naturalised epistemology is a misnomer, since it is not epistemology:

> But it is difficult to see how an "epistemology" that has been purged of normativity, one that lacks an appropriate normative concept of justification or evidence, can have anything to do with the concerns of traditional epistemology. And unless naturalized epistemology and classical epistemology share some of their central concerns, it's difficult to see how one could *replace* the other, or be a way (a better way) of doing the other. (Kim 1988, p. 391)

According to Kim, "for epistemology to go out of the business of justification is for it to go out of business." (Kim 1988, p. 391)

Kim's complaint about epistemology going out of business and Rorty's claim that epistemology would have nothing to do can be phrased differently: if methodology is a purely empirical study, then many philosophers of science will become unemployed. For example, in response to Giere's claim that, if epistemology is naturalised, philosophers of science would become theoretical scientists,[4] Siegel argues that there is no role for philosophers *qua* philosophers to play in such a naturalised version of philosophy of science, since these roles are already occupied by scientists:

> ...the suggestion that philosophers would be the theoreticians of the science of science will not do: in Giere's interdisciplinary naturalistic science of science, all the contributing disciplines have their theorists as well as their experimentalists. For example, the sociologists of science cited by Giere are theorists as much as experimentalists; similarly for the cognitive scientists and neuro-

[4]Siegel is responding to the following quote from Giere: "If the philosophy of science is naturalized, philosophers of science are on the same footing with historians, psychologists, sociologists, and others for whom the study of science is itself a scientific enterprise. The most philosophers of science could claim is to be the *theoreticians* of a developing science of science on the model of theoretical, as opposed to experimental, physics." (Giere 1985, p.343)

scientists Giere cites. (Siegel 1989 p. 373)

Kim asks what would be the new careers of the philosophers thus unemployed: "if normative epistemology is not a possible inquiry, why shouldn't the would-be epistemologist turn to, say, hydrodynamics or ornithology rather than [as Quine suggests] psychology?" (Kim 1988, p. 391)

Two responses might be made to this charge. The first is that both Rorty and Kim make the assumption that it is not possible to discuss questions of justification in a purely empirical approach, that is, that justification is an apriori notion. This, however, needs to be argued for. It may be the case that there is no apriori notion of justification to be had, while at the same time there may be an empirical account of justification (which is something philosophers may, if they wish, study). Indeed, this has been argued, and I shall address it in the next chapter. If this is the case, then it is true that naturalised epistemology is not apriori epistemology: but this is merely a dispute over terminology. Or it may be that there is no account of justification to be had at all (and philosophers really will face unemployment). A naturalist could respond that they have indeed changed the question, because there is no answer to the question: to use an analogy of Kornblith's, an atheist need not be concerned with theological questions of God's omnipotence, because the atheist believes that there is no God. In this case, epistemology would "go out of business," and with good reason: it would be a pointless pursuit. The second reply, however, is less attractive, since, as noted above, it does not serve as a justification for the purely empirical approach, but instead is an argument that all approaches are equally justified (i.e., not at all). Hence, in what follows, I shall pursue the first line of response.

2. Apriori justification is necessary for a naturalised philosophy of science

As I noted above, if the charge of changing the question is to have any force, it must be established that the questions not answered are necessary for the empirical approach. The charge of *begging* the question might be stated as follows: naturalised epistemology ignores (or cannot address) the issue of apriori justification, which must be settled if it is to succeed in its goal of giving a scientific account of how science works.

Siegel (1980) argues that not only is a switch to a purely descriptive account of scientific method unjustified, but that it is impossible. The goal of a purely empirical account of methodology is presumably an aposteriori theory of science. But, as Siegel points out, this project might well produce competing conflicting accounts of how scientists proceed (indeed, there are clearly an infinite number of competing methodologies to account for a given historical sequence). Such different accounts could not all be correct, because they conflict. A~~n~~ naturalist will therefore need to choose one account over another to succeed in explaining scientific activity in a purely empirical fashion:

> Quine says that "we are out simply to understand the link between observation and science." Well and good. But the link will be addressed by various theories, and we will have to discern between alternative accounts before we will have succeeded in "understanding the link." How we so discern will have to do with how we evaluate the various accounts; how we assess their claims. (Siegal 1980, p. 319)

Siegal seems to argue that the method employed to choose between conflicting theories of science must be apriori, which, naturally, an aposteriori account cannot provide. Siegel focuses on Quine's claim that naturalised epistemology is motivated by the same concerns as normative epistemology, namely, "understanding the link between observation and science." He claims that Quine equivocates between prescription and description, and that an empirical approach can only investigate the latter:

> The problem [with Quine's version of naturalism] is simply put: the notion of "understanding the link between observation and science" is equivocal. It equivocates between two distinct senses of "understanding the link": (a) understanding the psychological mechanisms by which scientific theories are produced, and (b) understanding the criteria by which we select one link over and against other links, one theory over and against other theories. This latter sense, of course, demands the evaluation of competing theories; the former does not. It is the former sense, though, that is amenable to empirical psychological research; the latter is not. (Siegal 1980, p. 319)

But the argument, as stated, assumes that there is no notion of justification available to advocates of a purely empirical approach. And indeed, this is Siegal's claim: that a purely empirical approach methodology can only give an account of all theories studied by the various sciences, that is, a purely empirical approach cannot distinguish between "good" and "bad" theories, or theories to be preferred and theories to be rejected:

> An account of the mechanisms of knowledge-acquisition, or the psychological processes involved in theory building, will account for the acquiring of theories we have rejected, as well as those we currently hold. Such an account, therefore, since it accounts for what we take to be "good" as well as what we take to be "bad" theories, cannot distinguish the good from the bad—which is to say, such an account cannot aid in justifying, or rationally preferring, one theory over another. (ibid)

If Siegal is correct, then a naturalised epistemology cannot succeed: to give an account of science is to produce a theory of how science works. But if an empirical approach cannot adjudicate between competing theories of scientific activity, then the empirical approach will not be able to complete its project of describing how science comes to its conclusions.

The response to Siegal's argument is that an empirical theory of science need not give an *apriori* account of which theories should be chosen over another, and that an aposteriori account will suffice. For example, perhaps the history of science, or some sociological theory of science, will demonstrate that a consensus is eventually reached on which competing theories are correct, without philosophical assistance, and that there will therefore be no difficulty in an empirical theory of science being chosen over its competitors. Thus the naturalist can respond to this argument with an empirical claim about convergence in science (this argument will be discussed in more detail below in section 4). Or, the naturalist could claim that "good" and "bad' can be adequately captured by empirical theories: indeed, it is to beg the question to assume that they can be defined only in an apriori manner. Hence, I conclude that this argument, like that of changing the question, is inconclusive.

3. Naturalism is self-refuting

One way of attempting to demonstrate that some notion of apriori justification is indispensable for epistemology is to argue that naturalism is self-refuting. If it could be shown that it is impossible to argue for a purely empirical approach without invoking some apriori notions, then if naturalism were true (i.e., if there were no apriori), this would be demonstrated by showing that naturalism is false (that is, naturalism begins from an apriorist point of view). It could be then concluded that naturalism cannot be correct. This objection is raised by Siegel and Almeder.

Siegel argues that Giere must argue from apriori premisses to his aposteriori position. Siegel's argument begins with the assertion that "there is no naturalistic explanation of scientific activity that would force, or even suggest" that it is a mistake to think that "the [normative, apriori] epistemology of science [is] an important dimension of philosophy of science." (Siegel 1989, p. 375) From this Siegel concludes that Giere must employ apriori arguments in favour of his position: "But if there is no naturalistic argument for Giere's conclusion that 'the study of science must itself be a science', then that conclusion is derived, illicitly on Giere's view, from non-naturalistic arguments. It appears that Giere must argue non-naturalistically [i.e., from apriori grounds] for his thesis that philosophy of science should be naturalized." (ibid) Thus, Giere cannot argue for the elimination of apriori arguments from epistemology, because in doing so, he invokes those very arguments:

> But since the dispute between the naturalist and the traditionalist [i.e., those in favour of apriori arguments] is itself a dispute within the philosophy of science, Giere can't argue that the *only* viable philosophy of science is naturalized philosophy of science. He needs non-naturalistic philosophy of science to argue the case for naturalism. (ibid)

As far as I can ascertain, Giere ignores this argument in his response to Siegel. But, as noted above, I believe Giere can be interpreted in two ways: first, he may be arguing apriori for his aposteriori approach by demonstrating, on apriori grounds, that such a programme cannot succeed. Second, he may be offering an aposteriori argument for his views. If the second interpretation is correct, then Sieg~~a~~el's argument is not a problem for Giere: Sieg~~a~~el's premise that "Giere is presumably *doing* philosophy of science [by this Sieg~~a~~el presumably means arguing from apriori premisses] when he argues that philosophy of science should 'go natural'..." (by which Siegel means that Laudan's "argument is non-natural [i.e., apriori]" ibid) is not true. Giere could reply that he is not arguing from apriori grounds: indeed he seems to argue that, empirically, apriori theorising has failed, while empirical investigations have not.

Even if the first interpretation of Giere is correct, Sieg~~a~~el's argument still does not show that Giere has refuted himself. Giere could argue that there is no contradiction involved in arguing within a programme to show that that programme cannot succeed—he may be employing a *reductio ad absurdum* argument as follows: if the apriori approach is true, it is false, therefore it is false. I will discuss this response more fully after considering a similar argument against Quine's naturalism made by Almeder.

Almeder levels the same criticism, that arguments for an empirical approach are self-refuting, but his main target is Quine. Almeder reads Quine's argument for naturalism as resting on two premises. The first is Hume's sceptical claims concerning aposteriori statements (i.e., that there is no certain knowledge. Quine's approval of Hume's scepticism can be found in his 1969 "Epistemology Naturalised"). The second premise is that there are no analytically true statements (Quine's arguments for this may be found in his 1953 [1980] "Two Dogmas of Empiricism"). These two premises imply that foundational attempts to establish certain knowledge of the world are doomed to failure: there are only synthetic statements, and these statements are uncertain, hence there are no certain statements. If all statements are synthetic, and if knowledge of synthetic statements cannot be certain, then certainty is not possible, and what Almeder (and Quine) call "first philosophy" is impossible. ("First philosophy" is perhaps in this context, a reference to Descartes's *Meditations on First Philosophy*, and refers, I think, to attempts to establish certain foundations for science. It could also refer to Aristotle's name for metaphysics, in Book V of the *Metaphysics*):

> ...as soon as we accept Quine's rejection of the analytic/synthetic distinction in favor of only synthetic propositions, Hume's argument casts a long despairing shadow over our ever being able to answer the sceptic because such propositions could never be certain anyway. The conclusion Quine draws from all this is that traditional epistemology is dead. There is no "first philosophy." (Almeder 1990, p. 264)

Almeder's objection to this line of argument is that he sees Quine as arguing that we should abandon apriori epistemology on apriori grounds: "[Quine] ...argues against there being a "first philosophy" by appealing to two premises both of which are sound only if philosophical arguments about the limits of human knowledge are permissible and sound." (Almeder 1990, p. 266) Almeder claims that the first premise, concerning skepticism, is not empirical: "The first premise consists in asserting that Hume's skepticism about factual knowledge is indeed established. Hume's thesis is certainly not empirically confirmable." (Almeder 1990, p. 266) From this Almeder seems to conclude that this premise cannot be legitimately argued for from purely empirical grounds: "The so-called "problem of induction" is a philosophical problem based upon a certain philosophical view about what is necessary for scientific knowledge. It is certainly not a problem in natural science or naturalized epistemology." (Almeder 1990, pp. 266-7) Similarly, argues Almeder, the rejection of the analytic/synthetic distinction is a non-empirical matter: "The second premise is the denial of the analytic/synthetic distinction; and that is a thesis largely resting on a philosophical argument about the nature of meaning." (Almeder 1990, pp. 267)

Almeder concludes that since Quine relies on non-empirical arguments to support his naturalistic approach, Quine must assume that at least some apriori propositions are valid: "Such premises only make sense within a commitment to the validity of some form of first philosophy and the legitimacy of traditional epistemology." (Almeder 1990, p. 267). This, Almeder asserts, makes Quine's thesis self-refuting, since "philosophical" (presumably this means "apriori") arguments cannot support an argument for the elimination of the "philosophical": "...it is obvious that Quine's argument itself for naturalized epistemology is a philosophical argument, which, *ex hypothesi*, should not count by way of providing evidence for the thesis of naturalized epistemology." (1990, p. 266)

I do not believe that Almeder has established that Quine's arguments for a purely empirical approach to philosophy (as standardly interpreted) are self-refuting. If we grant Almeder's contention that Quine argues from apriori grounds, then Quine's argument can be construed as follows:

(1) All statements are either analytic or synthetic
(2) There are no analytically true statements.
(3) All synthetic statements are uncertain.
Therefore,
(3) All statements are uncertain.

This is a valid argument. And, if all statements are uncertain, then there is no such thing as first philosophy, i.e., a set of statements whose truth is known independently of any empirical data. By substitution

(3') There are no statements whose truth is known independently of experience.

The question then arises as to how we know the truth (3'). If it follows from statements known independently of experience, then it seems that the truth of (3') is known independently of experience. But this means that if (3') is true, then it is false, and so it is. This is what I take to be the point of Almeder's argument, and I agree that it is forceful with respect to naturalism when it is defended from apriori grounds: it shows that such an approach would indeed be self-refuting..

However, it may also be concluded for the above considerations that the truth of (3') must be determined by empirical means. If this is the case, then there is no problem with self-refutation. This is the approach that I believe that Quine would take. Quine's argument in "Two Dogmas" against the analytic/synthetic distinction is that all attempts so far to draw the distinction in an apriori manner have been unsuccessful, but that we can describe something like the distinction in empirical terms (as an hypothesis about the use of language). Quine's arguments (in "Epistemology Naturalized") for the thesis that empirical statements can never be certain are also negative in this manner—Quine argues that all attempts to provide a means of certifying empirical statements have failed.

The question then arises as to how the purely empirical approach can be justified on aposteriori grounds. One answer to this question is that the empirical approach is in some sense self-justifying. I shall examine this in the next section.

4. The Circle Argument

I shall pose a dilemma: Naturalism may be defended on either non-empirical or empirical grounds; if the first, then it is self refuting (as I argued above). In this section I argue that if the second, then arguments for naturalism are circular. The circle argument runs as follows: If we attempt to argue in an empirical manner for the adoption of a purely empirical approach, then we have assumed what we are trying to establish, that the empirical approach is correct, by assuming that it is correct. One way around the problem of arguing in a circle (and also, as I argued in the previous section, of self-refutation) would be to hold that scientific statements, and hence a purely empirical approach to philosophy, are self-justifying. That is, naturalism needs no further justification. The response to such an argument will establish if naturalism is circularly justified or not: if it is possible to establish empirical arguments in a self-justifying manner, then the circle argument is not a difficulty for naturalism. In this section I shall first examine the so-called circle argument against naturalised epistemology. I shall then turn to Quine's attempts to deny the need for philosophical argument in favour of naturalised epistemology. I then turn to Giere's attempts to block the circle argument.

Quine's response to the criticism that he is arguing in a circle, justifying an empirical approach to methodology by assuming that the empirical approach is valid, is that the only need for justification of the empirical approach comes from a need to respond to sceptical challenges about the veracity of empirical evidence. These challenges, Quine claims, are rightly answered in an empirical manner. Quine characterises the sceptic as arguing from scientific grounds, using parts of science (in particular, psychology) to pose difficulties for the foundations of science. Quine responds that since the difficulties raised by the sceptic are empirical, it is not circular to appeal to scientific findings in justifying an empirical approach. I will consider the theses that sceptics rely on (parts of) science, and that this justifies the use of scientific data in refuting the sceptic in turn.

First, for Quine, "sceptical doubts are scientific doubts." (Quine 1975, p. 68) Quine takes as an example "ancient skepticism", which he claims "challenged science from within" by citing "familiar illusions to show the fallibility of the senses..." But, according to Quine, the sceptic had to admit that there was an external world in order to be able to claim that the way we perceived the world may be delusive: "...this concept of illusion itself rested on natural science, since the quality of illusion consisted simply in deviation from external scientific reality." (Quine 1973, p. 3) So for Quine, "scepticism is an offshoot of science. The basis for scepticism is the awareness of illusion, the discovery that we must not always believe our eyes." (Quine 1975, p. 67) Quine characterises the modern version of scepticism as one which argues from the paucity of our sensory capabilities to the underdetermination of our theories by empirical evidence:

> Science itself teaches that there is no clairvoyance; that the only information that can reach our sensory surfaces from external objects must be limited to two-dimensional optical projections and various impacts of airwaves on the eardrums and some gaseous reactions in the nasal passages and a few kindred odds and ends. How, the challenge proceeds, could one hope to find out about that external world from such meager traces? In short, if our science were true, how could we know it? (Quine 1973, p. 2)

Quine seems to conclude from the claim that "sceptical doubts are scientific doubts" that if one wishes to question the truth of science claims, one must do so from scientific grounds.

This leads to Quine's second thesis, that the sceptic's reliance on the notion of illusion implies that scientific evidence may be used as defence against the sceptic. That is, since the sceptic uses (parts of) science, so may the defender of science use (all of) science:

> It was science that demonstrated the limitedness of the evidence for science. And it would have befitted the epistemologist, then as now, to make free use of science in his effort to determine how man could make the most of those limited sources. (ibid, p. 2-3)

Clearly, in confronting this challenge, the epistemologist may make free use of

> all scientific theory. His problem is that of finding ways, in keeping with natural science, whereby the human animal can have projected this same science from the sensory information that could reach him according to this science. (ibid, p. 2)
>
> This fear of circularity is a case of needless logical timidity, even granted the project of substantiating our knowledge of the external world. The crucial logical point is that the epistemologist is confronting a challenge to natural science that arises from within natural science. (ibid)

In particular, Quine claims that to challenge the veracity of scientific knowledge, the sceptic must assume that there are "bodies": "Illusions are illusions only relative to a prior acceptance of genuine bodies with which to contrast them. In a world of immediate sense data with no bodies posited and no questions asked, a distinction between reality and illusion would have no place." (Quine 1975, p. 67) But, according to Quine, to postulate "bodies" and their "regularities" is to accept "rudimentary physical science":

> The positing of bodies is already rudimentary physical science; and it is only after that stage that the skeptic's invidious distinctions can make sense. Bodies have to be posited before there can be a motive, however tenuous, for acquiescing in a non-committal world of the immediate given.
>
> Rudimentary physical science, that is, common sense about bodies, is thus needed as a springboard for skepticism. It contributes the needed notion of a distinction between reality and illusion, and that is not all. It also discerns regularities of bodily behaviour which are indispensable to that distinction. The skeptic's example of the seemingly bent stick owes its force to our knowledge that sticks do not bend by immersion; and his examples of mirages, after-images, dreams, and the rest are similarly parasitic upon positive science, however primitive. (Quine 1975, p. 67-8)

Quine's account of scepticism provides a model of his view of science. Quine holds that there are two ways in which science describes our relation with the external world. First, physics (together with a scientific description of our sensory mechanisms) describes the components of the world and the components of our bodies, and their interactions. Secondly, psychology then tells us how we translate these interactions into theories concerning the world. If we assume the relevant portion of physics to be correct, then we have the task of explaining how we come to know it to be correct, which is the second, psychological, task. If we assume our psychological translations of data to be correct, then we must explain what the source of the data must be like, which is the second task, that is, one for physics. According to Quine, we should examine subjects in given environments, and study how they respond. We are to give the subject "a certain experimentally controlled input—certain patterns of irradiation in assorted frequencies, for instance..." (Quine 1969, pp. 82-83), and then study what "the subject delivers as output", which is taken as "a description of the three-dimensional world and its history." We are then, Quine seems to claim, in a position to account for our knowledge, by putting ourselves in the position of the experimental subject:

> We are studying how the human subject of our study posits bodies and projects his physics from his data, and we appreciate that our position in the world is just like his. Our very epistemological enterprise, therefore, and the psychology wherein it is a component chapter, and the whole of natural science wherein psychology is a component book—all this is our own construction or projection from stimulations like those we were meting out to our epistemological subject. (ibid, p. 83)

Thus, according to Quine, it is possible to account for knowledge in a purely empirical fashion: we need only show how people in fact learn from experience. We can then proceed to learn from experience in the same way.

According to Quine, this is not a circular argument for naturalism, that is, it is not assuming that our scientific theories are correct to show that they are correct, for the following reason: we need to justify an aposteriori approach to epistemology only in response to someone who does not believe such an approach can lead to knowledge (the sceptic). But the only way to show that the aposteriori approach does not lead to knowledge is to assume a part of a scientific theory (and hence is not an apriori matter, but an aposteriori one). But the part of the scientific theory which the sceptic assumes is sufficient for a psychological account of learning, thus answering the sceptic's doubt about the possibility of knowing anything about the world. For example, the sceptic who uses the example of sticks looking bent when placed in water to question the veracity of our claims about sticks in water assumes something about sticks, water and perception. According to Quine, the way to respond to the sceptic is to show how these assumptions confirm our theories about the behaviour of sticks immersed in water. We need not, says Quine, attempt to show from a position of complete doubt how to gain knowledge of the external world: "Cartesian doubt is not the way to begin." The way to ~~being~~ begin epistemology, for Quine, is to develop a theory of learning: "Epistemology is best looked upon, then, as an enterprise within natural science... Retaining our present beliefs about nature, we can still ask how we can have arrived at them." (Quine 1975, p. 68).

If this is Quine's response, I do not think that it is adequate. In particular, Quine's characterisation of the sceptic as assuming parts of science to refute science does not seem accurate. Barry Stroud argues that the sceptic need only demonstrate that the data we have are compatible with many hypotheses—a hypothesis that we are only dreaming, that we are being deceived by an evil demon hypothesis, that we are a brain in a vat, or that there are in fact physical objects—in order to argue that an empirical approach is not justified by empirical data. To use Stroud's example, it is the same situation as looking into our garden and seeing a bird which could either be a canary, a goldfinch or a goldcrest. Without any independent reason for believing that the bird was one or the other, we would not know if the bird was a goldfinch or not; that is, we would not know the truth about the goldfinch hypothesis. Similarly, the sceptic argues that we do not have knowledge about the world, because many hypotheses are consistent with our data:

> It [scepticism] simply says that none of the competing 'hypotheses' about what is true beyond the data can be known to be true; in fact, that we can have no more reason for believing any one of them rather than others on the basis of the only sensory data we can ever have. If our data are so inevitably restricted in relation to what we claim to know on the basis of them, the conclusion that we can know nothing beyond the data is no over-reaction at all. (Stroud 1984, p. 233)

Quine's response to Stroud's argument is difficult to understand, since it seems to miss the point: he responds that other theories may indeed consistently explain the data, but the theories we use and regard as true have made correct predictions, so far. But if scientific hypotheses ceased to be successful, then, according to Quine, we could choose to construct hypotheses to explain the data based on, say, dreams. However, Quine claims, this would still be an aposteriori approach to epistemology:

> Experience might, tomorrow, take a turn that would justify the skeptic's doubts about external objects. Our success in predicting observations might fall off sharply, and concomitantly with this we might begin to be somewhat successful in basing predictions upon dreams or reveries. At that point we might reasonably doubt our theory of nature in even its broadest outlines. But our doubts would still be immanent, and of a piece with the scientific endeavour. (Quine 1981, p. 475)

Quine's response is puzzling, since it is does not address the objection that there are infinitely many hypotheses which are compatible with any given data. His response seems to be that our theories of the world are well confirmed, and the sceptic's are not:

> *True, we must hedge the perhaps too stringent connotations of the verb 'know'; but such is fallibilism.*
>
> Thus, in keeping with my naturalism, I am reasoning within the overall scientific system rather than somehow above or beyond it... *The skeptic repudiates science because it is vulnerable to illusion on its own showing; and my only criticism of the skeptic is that he is overreacting.* (italics mine, ibid, p. 474-475)

To this Stroud responds, correctly, I believe, that Quine misses the point: the sceptic does not argue that there is a different scientific theory of knowledge which is better confirmed by some data, but that there are hypotheses ~~concerning~~ about our relations with the external world which are equally confirmed by the data. It is not that our scientific hypotheses may be refuted by experience, and dream hypotheses confirmed, but that the two will always be compatible with the data.

It thus seems that Quine is not correct about the need for the justification of an empirical approach to epistemology. The sceptical challenge to naturalism may be construed as aposteriori grounds: the sceptic argues that there are (and will always be) equally many hypotheses compatible with the data. This is correct, and so it does not seem possible to give an aposteriori response. I conclude that the circle argument does indeed raise a difficulty for naturalism: if epistemology is to be an empirical science, then it may be asked how we know that the theories that science produces are correct. An empirical response that one empirical theory of knowledge is compatible with the data, while other, competing theories, are not correct, since there will always other theories compatible with the same data. Before I conclude that if naturalism is argued for from aposteriori grounds it is argued for in a circular manner I must examine Giere's response this argument.

Giere claims that his argument for naturalism is 'virtuously', and not viciously, circular. I shall try to explicate this claim. Giere argues that since humans have evolved certain capacities (e.g., to survive) and it is with these capacities that we made science and interpret the external world. Giere claims that it is wrong to start from "Cartesian doubt," since no human has ever been in a situation of complete doubt concerning the external world:

> The evolution of humans required some abilities to construct rough representations of the world which make it possible to deal at least moderately effectively with middle sized objects, including other humans... These skills comprised the evolutionary heritage with which humans began their construction of society and, eventually, of modern science. Thus, no human ever faced the world in a state of Cartesian doubt. Why, then, should anyone think that our understanding of science requires a notion of "justification" that rejects any appeal to even the most rudimentary empirical knowledge? (Giere 1989, p. 379)

Giere, like Quine, seems to believe that one need to justify the empirical approach to answer the sceptic. In particular, he seems to hold that apriori epistemology requires that we begin by doubting all knowledge of the external world (this is Descartes' method). Giere objects that even if we follow Descartes through the *Meditations*, we will carry our skills aquired by evolution with us. Thus, even a position which aims to begin by assuming nothing about the external world will fail in doing so, because the arguments must include something obtained empirically. So, for Giere, since we must use empirical methods, we should use them in a way which, he claims, will improve those methods: we can examine the ways in which our intellectual capacities are restricted by our evolutionary heritage, and then apply the results of this examination to our epistemology. Giere likens this to a positive feedback loop:

> By looking back at evolutionary history, scientists themselves can better understand their own cognitive situation and investigate the development of their own cognitive capacities. What seem to the traditional epistemologist like vicious circles are, in this alternative picture, "positive feedback loops." Using our evolved cognitive capacities, we extend our knowledge of the world, including our knowledge of our own cognitive abilities. This latter knowledge helps us to extend our knowledge of the world still further. And so on. The existence of these loops is not a limitation that must be overcome by some special form of philosophical analysis. On the contrary, it is one of the things that makes modern science so powerful. (Giere 1988, p. 13)

Thus, according to Giere, the circle argument is an argument for the empirical approach to methodology. For example, the epistemologist could concentrate on what is taken to be good scientific work. As the definitions of "good scientific work" become clearer, science will become more powerful, thus allowing better definitions of "good scientific work." This process of refining the meaning of terms is common in science. The notion of heat originally included spices as heat producing objects, since they made the tongue warm. But, as science got better theories of heat, spices were no longer classed as heat producing objects.

Giere's argument is much the same as Quine's: he seems to believe that the only reason one might not accept a purely aposteriori approach to methodology is because it is not well confirmed. But, as I pointed out regarding Quine's response to Stroud, the difficulty is not that sceptical hypotheses are or are not well confirmed, but that they are equally compatible with the data. Giere likens the relation of empirical theories about the external world to the external world to a positive feedback loop. But to have a positive feedback loop, we must be able to show that evolution has in fact correctly tailored our perceptions to the world. Systems which use positive feedback, such as steam engines, or certain electronic circuits, use their own output to regulate the values of their input, which is in turn used to regulate the output. The aim of such a system is to keep the output and the input within a certain range. Sometimes systems display 'negative feedback' loops, such as the 'howl' in a poorly adjusted public address system. In other words, the data have to be compatible the theories. One who questions the aposteriori approach may be seen as arguing that many other theories can be in a positive feedback relationship with empirical data, and ask why we should choose any particular one over the other.

An advocate of the purely empirical approach might suggest that science can choose one particular confirmation theory on empirical grounds. This however, begs the question of how science could lead us to make such a choice. Presumably, the justification of a particular choice on empirical grounds would also stand in need of justification. This justification may be apriori or aposteriori. If it is apriori, then this shows that there can be apriori justifications, and that aposteriorism is false. If it is aposteriori, then the argument is circular. It would be possible to appeal once again to aposteriori grounds in order to avoid the circle, but this would lead to an infinite regress. The problem with accepting an infinite regress is illustrated by the venerable joke about a man who prefers a broken watch to a watch which runs slowly. His reason is that the broken watch is perfectly accurate twice a day. Obviously, the broken watch is worthless, since he does not know when it is accurate. Similarly, if it cannot be shown when an empirical method will accurately identify a correct method, the empirical method is similarly worthless.

I conclude that Giere's response to the circle argument, like Quine's, is inadequate. Hence I conclude that if one argues for an empirical approach to epistemology from aposteriori grounds, one is arguing in a circle. In particular, an aposteriori argument for naturalism requires that sceptical hypotheses be dismissed out of hand, in favour of the more familiar empirical hypotheses. But this is to assume that an empirical approach is correct. Such an argument for naturalism is as unenlightening as it is unconvincing. In the next section I shall argue that this is one of the horns of a dilemma advocates of naturalism must face.

C. Conclusion

There are many arguments against naturalism, but I only find two convincing. These two form a dilemma, and this dilemma shows that an empirical approach to epistemology cannot succeed. Naturalism may be argued for from apriori or aposteriori grounds. In section B3 I argued that if an aposteriori approach to science is argued for in an apriori manner, it is self-refuting: the arguments would show that we can come to knowledge by apriori means, and thus, to some extent, justify an apriori approach. In section B4 I argued that if an aposteriori approach is argued for in an aposteriori manner is argued for circularly. I conclude that the aposteriori approach is either wrong (self-refuting) or unjustified (argued for in a circle). One could reject the notion of justification (by accepting circular arguments), but this would entail that any approach to methodology is no better than another, and, *a forteriori*, that an empirical approach is no better than an apriori approach.

However, in the first part of this thesis I concluded that apriori approaches are also not justified. This motivates the search for a compromise. In the next section I shall examine in detail Quine's account of methodology to determine if it offers such a compromise. I shall conclude that it does not, but provides the grounds for developing such a compromise.

CHAPTER 5: Quine's Naturalism: A new reading of "Epistemology Naturalized"

> It is reconciling what Quine says here with what Quine says elsewhere that is difficult and confusing. I am not claiming that it is impossible however; a lot, if not all, of what Quine says can be reconciled. What I claim is that Quine's position is much more complicated than is generally realized. (Putnam 1982, p. 296)

> Locke, Hume, and, towards the end of the century, Bentham and the philosophical radicals agreed in denying the existence of any such faculty as an intellectual intuition into the real nature of things. No faculty other than the familiar physical senses could provide that initial empirical information on which all other knowledge of the world is ultimately founded. Since all information was conveyed by the senses, reason could not be an independent source of knowledge, and was responsible only for arranging, classifying and fitting together such information, and drawing deductions from it, operating upon material obtained without its aid. (Berlin 1939, pp. 40-41)

A. The conceptual and doctrinal programmes
B. Quine's history of the conceptual and doctrinal programmes from Descartes to Russell
C. Carnap's *Aufbau*
D. The failure of the apriori conceptual programme
- **1. The indeterminacy of translation**
- **2. Against confirmational holism**

E. (Part of) Epistemology naturalised
F. An example of the conceptual programme naturalised—the observation sentence
- **1. The method of radical translation**
- **2. Observation sentences**

G. Quine on the doctrinal programme
- **1. Canine induction**
- **2. Infant induction**
- **3. Scientific learning**
- **4. The doctrinal programme conventionalised**
- **5. Quine's epistemology defended**

H. Quine's programmes and the dilemma of naturalism
I. Quine on the unity of science
J. Quine's confirmation theory
K. Conclusion

As we have seen, the naturalist position that is commonly attributed to Quine faces severe difficulties. In fact, it seems to be indefensible, which leads one to wonder how commentators can accept with equanimity the idea that so eminent a philosopher as Quine adheres to it. This chapter is a defence of Putnam's statement, quoted above, that reconciling Quine's writings on epistemology is indeed "difficult and confusing." I shall argue that the best reading of Quine is not the one normally ascribed to him and described in the previous chapter, but a more complicated empirical-cum-conventional account of epistemology, much along the lines as those described by Berlin above. I shall argue that many authors have misunderstood Quine's views, and that they may have done this because they emphasised too much his 1969 "Epistemology Naturalized." This paper is, I believe, only an outline of Quine's views, which should be filled in by Quine's other writings.

I will now document that Quine is in fact widely interpreted as championing a purely empirical approach devoid of any normative considerations to justify my claim to a new reading of "Epistemology Naturalized." Kim claims that "Quine is urging us to replace a normative theory of cognition with a descriptive science" (Kim 1988, p. 389); James Maffie tells us that "Quine... represents the locus classicus of eliminativism." Thus according to Maffie, Quine's epistemology is one in which "[t]he normative dimension of knowledge is abandoned; stock-in-trade normative notions like justification, reasons, and evidence play no role in their [the eliminativist's] epistemologies." (Maffie, 1990, pp. 284-285) Hilary Kornblith takes the same position. (Kornblith 1985, p. 3-4) Richard Creath says that Quine "conceives of epistemology as a branch of empirical psychology and hence as a descriptive enterprise rather than as a normative one." (Creath 1990, p. 41) Harvey Siegal (1980, pp. 317-319) agrees with this interpretation, as does Robert Almeder (1988, p. 263).

Unlike those who endorse the standard view, I do not believe that Quine deliberately set aside (or overlooked) normative or justificatory considerations. Indeed, I think that such considerations are central to his epistemology. Certainly there is an empiricist strain in his philosophy, but I believe that this leads Quine to argue for the naturalisation only of a part of epistemology, namely what he calls the conceptual programme (the programme to clarify epistemological concepts). The "doctrinal programme" (the programme of the justification of epistemological concepts) is for Quine, in my understanding, not to be purely empirical, but partly empirical and partly conventional. Thus Quine, unlike most of the authors quoted in the preceding two paragraphs, does not identify 'normative' with 'apriori.' Quine's writings on the status of justificatory concepts are indeed confusing: this necessitates a careful reading in order to locate the sources of this confusion.

I interpret Quine's views on epistemology as follows. Since Descartes, according to Quine, epistemologists have undertaken two tasks—to clarify basic epistemic concepts, and then to justify knowledge claims formulated in terms of the clarified epistemic concepts. Quine refers to the former as the "conceptual programme," the latter as the "doctrinal programme." I examine this dichotomy in section A of this chapter. Quine clearly sympathises with the empiricist tradition in epistemology, as he concentrates on such philosophers as Hume, Mill, Bentham, Russell and Carnap (and does not mention Kant in this context). And, I shall argue in this chapter, Quine sees himself as continuing this empiricist tradition.

Quine sees the doctrinal programme, insofar as it is concerned with deductive justification of knowledge, i.e., the attempt to ground knowledge of the external world by deduction from indubitable foundations, as a failure. He concentrates instead on the conceptual programme, which he sees as culminated in Carnap's *Der Logische Aufbau der Welt* (this takes us through section B). In this work, Carnap attempted to complete the conceptual programme by defining scientific concepts solely in terms of a single subjective predicate ("recollection of similarity") and the logic of *Principia Mathematica*. This project is described in section C.

Quine sees Carnap's attempt as a failure, but an enlightening one: Carnap failed, in Quine's view, because his reductionist strategy employed in the *Aufbau* failed; indeed, Quine argues that it is impossible to reduce scientific theories to logical and sense-data terms. This is his famous "indeterminacy of translation thesis," which is based on a verificationist theory of meaning and on the Duhem thesis that "isolated" theories are neither confirmable nor falsifiable. I shall argue that the indeterminacy of translation, if there is such a thing, does not show that reduction is impossible. Quine's argument does make clear, however, his views on so-called meaning holism, the view that only large portions of scientific theories have meaning. I discuss these issues in section D.

Having determined to his satisfaction the reasons why the conceptual programme failed, Quine presents a positive account of how epistemology should proceed, while avoiding the failures of the previous attempts. According to Quine, what is required is an empirical approach to epistemology (section E). Quine's positive account is ambiguous, however, since his arguments only seem to entail naturalisation of the conceptual programme, and not the doctrinal (section F). To try to resolve this ambiguity, I examine Quine's proposals for an empirical conceptual programme, which are adumbrated in "Epistemology Naturalized," and laid out more fully in *Word and Object* and *The Roots of Reference*. This new Quinean programme is based on a behaviourist account of language learning, which, Quine seems to think, is required by empiricism. He gives the account in two forms: how a translator learns a new language, and how an infant learns to use language. According to Quine, translators and infants learn certain basic statements, understood by all speakers of a language, by a process of induction. These statements are what Quine calls observation sentences, and it is to them that Quine looks for new foundations for both the conceptual and the doctrinal programmes. I examine the former in section G. Quine claims that a further study of language learning will reveal the processes by which we learn other parts of language, and how we learn scientific theories. This, Quine claims, will reveal the process by which science proceeds. (section H)

Quine sees his version of the conceptual programme as continuing in the empiricist tradition: clarity, the aim of the conceptual programme, is conceived of solely in behaviourist terms. How Quine views his empiricist programme's relation to the traditional doctrinal programme is not, however, clear. I shall argue that for Quine induction is used by all humans, but its epistemic status is conventional. Nonetheless, Quine seems to believe that we do not in fact choose between conventions, but that we all accept by our very use of language what Quine merely calls "science." So, the grounds we employ for justifying theories are conventional and the choice of a convention is for Quine unproblematic, because he claims that there is universal agreement on the basic conventions. In section I, I shall argue that for Quine these basic conventions are a mixture of logical, subjectivist and frequentist accounts of probability.

There is much to disagree with in Quine's version of epistemology: his argument for the indeterminacy of translation, his claim that this accounts for the failure of reductionism, his verificationist criterion of meaning, his holism, his theory of translation, and his behaviourist view of animal learning. I shall not criticise these at length, but will concentrate instead on his conventionalism. While I wish to argue in the final chapter for a type of conventionalism, I shall, in the final section of this chapter dispute Quine's claim of universal agreement about inductive methods.

A. The conceptual and doctrinal programmes

As noted above, Quine conceives epistemology as "concerned with the foundations of science" (Quine 1969, p. 69) and foundational studies as divided into two types of activity—conceptual and doctrinal, corresponding, respectively, to the analysis and the justification of epistemological concepts: "conceptual studies are concerned with meaning," "clarifying concepts by defining them, some in terms of the others", and with "explaining the notion of body in sensory terms," while the doctrinal programme is concerned with "truth", "establishing laws by proving them, some on the basis of others" and "justifying our knowledge of truths of nature in sensory terms." (Quine 1969, pp. 69-71) (Quine presumably means that the conceptual programme is concerned not only with explicating the notion of physical objects in terms of sense data, but also with the relations which can hold between those bodies: otherwise it could not account for, e.g., gravity.) Quine never repudiates these twin epistemological tasks, which seem to be a modern version of the Cartesian programme of the *Meditations* (indeed, Quine refers to the doctrinal programme as the "Cartesian quest for certainty" (ibid, p. 74)). According to Quine, the aim of epistemology on the conceptual side is to define less clear concepts in terms of more clear concepts; and on the doctrinal side to prove less obviously true laws from ones that are more obviously true, where the basic defining concepts will be "clear and distinct" and the basic laws will be "self-evident truths." (The last two quoted phrases are Quine's, ibid, p. 70) The appeal of this twin approach to epistemology is that if the ideal is realised we shall be able to translate obscure concepts into clear ones, which can then be used to state our laws. The truth of some of these laws will be obvious, and that of others will be seen in their derivation from them.

Quine illustrates this division of tasks in foundational epistemological studies with mathematics, since epistemology "include[s] the study of the foundations of mathematics as one of its departments." (1969, p. 69) In the history of the epistemology of mathematics, an attempt was made to reduce mathematics to the seemingly self-evident principles of logic. This, Quine points out, proved not to be possible. It turned out that mathematics can at best be reduced to logic and set theory, which Quine says "is a disappointment epistemologically, since the firmness and obviousness that we associate with logic cannot be claimed for set theory." (ibid, p. 70) And he also points to Gödel's result that any part of mathematics which contains arithmetic cannot be consistently axiomatised, which precludes an encapsulation of mathematics in a set of (self-evident or otherwise) axioms. Thus, we can complete neither the conceptual nor the doctrinal programmes for mathematics. However, the reduction of mathematics to set theory does, in Quine's view "enhance clarity, but only because of the interrelations that emerge and not because the end terms of the analysis are clearer than others." (ibid) But this extra clarity gained by the reduction of mathematics to logic and set theory does not serve the doctrinal aims of the Cartesian programme, since it does not show that there is an epistemologically privileged set of statements: "Reduction... does not do what the epistemologist would like of it: it does not reveal the ground of mathematical knowledge, it does not show how mathematical certainty is possible." (ibid) Quine holds that the same problems which arose for the conceptual and doctrinal programmes in the foundations of mathematics also plague epistemology generally. I will spend the next section exploring this parallel.

B. Quine's history of the conceptual and doctrinal programmes from Descartes to Russell

Quine believes that the dual doctrinal/conceptual approach has been the way in which epistemology has been investigated at least since Descartes. To support this claim, Quine provides a thumbnail history of epistemology from Hume to Carnap. In this section, I only consider Quine's version of this history from Hume to Russell. (Quine gives a similar account in his 1953 [1980] "Two Dogmas of Empiricism", pp. 38-41). Although the division between conceptual and doctrinal programmes in epistemology is clearly Cartesian, Quine evidently sees Descartes' attempts to complete the doctrinal and conceptual programmes as failures, since his history begins with David Hume. Hume, Quine tells us, "identified bodies outright with sense impressions." (Quine 1969, p. 71) On this analysis, the monitor on my desk changes as I type in new characters, and so turns into a different body. As my sense impressions change, so do the bodies. Quine sees this identification of physical objects with sense impressions as a beginning for the conceptual programme in that it gives a clear analysis of the notion of a physical object. However, Hume's analysis also has the counter-intuitive consequence that bodies do not persist over time or through physical displacement.

Quine also points out that Hume's analysis of physical objects yields no certainty concerning general or future statements about objects, and therefore does not (substantially) advance the doctrinal programme. We can be certain that we are experiencing particular impressions, and so, Quine believes, singular statements about bodies (i.e., quantifier-free statements) can be certain. If the statement 'A monitor is in front of me' is translated as 'A glowing box-shaped object is here, now', we can be certain that a monitor is in front of us. "But," according to Quine, "general statements, also singular statements about the future, gained no increment of certainty by being construed as about impressions." (Quine 1969, p. 72) For example, the statement that all monitors must be connected to a power source to glow is in no way made certain by this analysis. From this Quine concludes that basing the conceptual programme solely on current sense data cannot assist the doctrinal programme. Quine apparently sees no progress concerning the doctrinal programme after Hume: his epigram "The Humean predicament is the human predicament" (Quine 1969, p. 72) seems to indicate that Quine thinks the doctrinal programme hopeless. But this means that the doctrinal programme, which it will be recalled is concerned with "establishing laws by proving them, some on the basis of others", cannot be fulfilled. The word "proving" is central to the description of the doctrinal programme. If it is to mean "deducing," then it is surely the case that the doctrinal programme cannot be completed. I believe that Quine is in fact referring to deduction, and it is this type of foundational programme that stopped at Hume. I shall argue (beginning in section G) that Quine envisages another type of foundational programme which does not aim to demonstrate the absolute certainty of some statements, but rather their probability.

Although Quine views the doctrinal programme pessimistically, he does believe that the conceptual programme has made some progress since Hume. Quine credits the main advance to Bentham and his notion of contextual definition or paraphrasis, which was developed and formalised in Russell's Theory of Descriptions. Quine claims that this theory is an advance because it allows terms denoting physical objects persisting over time to be translated into statements containing only sense data terms. Russell's developed his theory of descriptions in order to eliminate non-referring singular terms. These terms were a problem for his theory of meaning, as I shall show. His solution provided a general method, later exploited by Carnap, to eliminate singular terms referring to physical objects.

If we hold that the meaning of a proper name is the object to which it refers, then we encounter well-known difficulties with such names as "unicorn," or, to use Quine's example, "The round square cupola on Berkeley College" (Quine1953a [1980], p. 6) It seems that we must choose to classify sentences containing terms that do not refer such as "The round square cupola on Berkeley College is pink" as meaningless, since they contain a meaningless term, or admit that the sentence has meaning, and hence admit the existence of round square cupolas. However, while the sentence seems to be meaningful, there are no such things as round square objects.

Russell attempted to solve the difficulty of meaningful sentences containing non-referring names by distinguishing between logically proper names (those that refer) and terms which are grammatically terms, but do not refer. Non-referring names are analysed in Russell's theory so that, while they have no meaning in isolation, they do gain meaning when they are part of a sentence. According to Russell, we should translate the sentence "The round square cupola on Berkeley College is pink" into (first order) logic as

$$\exists y[\forall x(\text{Round}(x) \wedge \text{Square}(x) \wedge \text{Cupola}(x) \wedge \text{On Berkeley College}(x) \leftrightarrow x=y) \wedge (\text{Pink}(y))]$$

More generally any sentence of the form "the x such that f," can be translated as

$$\exists y[\forall x(fx \leftrightarrow x=y) \wedge (\ldots y\ldots)]$$

where fx is a propositional function. The left hand side of the main conjunct expresses the uniqueness implied by 'the', the right hand side completes the sentence. Russell's notation for the definite description was '(...(ιx)fx...)' (Because of the limitations of the modern word processor, I use an iota where Russell used an inverted iota).

A definite description can be used as the argument of a function or the value of a predicate, since definite descriptions can be used in place of constants. Hence we may write instead of Ga, $G(\iota x F(x))$, where $\iota x F(x)$ is the definite description of a. However, as can be seen from the above example, a definite description cannot be substituted into a sentence: rather, Russell's theory gives us a way to transform a sentence not containing a definite description into a sentence with a definite description. Thus, Russell used the lacunae, '(...(ιx)fx...)' and '...y...', because in his theory, a definite description only has meaning in context, hence the name "contextual definition." Russell referred to definite descriptions as "incomplete symbols."

According to the Theory of Descriptions, troublesome sentences about unicorns and round square objects do not contain names, but variables, and so do not refer to any particular object: they contain instead definite descriptions which, grammatically speaking, act like proper names, but are in fact not. Definite descriptions differ from proper names in that they only have meaning in context. Thus, Russell claimed to have shown how to talk about nonexistent entities without referring to them (that is, without using proper names).

Russell not only gave a clear analysis of troublesome terms such as "the round square," he also, according to Quine, advanced the conceptual programme by analysing physical objects in terms of definite descriptions. Russell's proposal was to identify bodies with sense data by means of definite descriptions: "The 'thing' of common sense may in fact be identified with the whole class of its appearances." (Russell, "The relation of sense-data to physics", 1917, p. 115) If we wish to talk about physical objects (or "bodies" in Quine's terminology), according to Russell, we need only give a definite description of that thing as a set of events through time: we do not need to refer to something 'in itself.' This is clearly an advance over Hume's definition of a physical object, since we can now give meaning to the ordinary use of object terms, that is, as terms referring to objects persisting over time, without actually assuming that an object is anything more than a collection of phenomena: we do not need to use proper names to refer to things. As we have seen, we could eliminate the names of specific objects in favour of definite descriptions of objects by associating a variable with certain properties: thus, a body can be characterised fully in terms of a variable for which a class of sensory impressions holds. This means, according to Quine, that we could make "sense of talk of bodies, even granted that impressions were the only reality. One could undertake to explain talk of bodies in terms of talk of impressions by translating one's whole sentences about bodies into whole sentences about impressions, without equating the bodies themselves to anything at all." (Quine 1969, p. 72)

The completed conceptual project would have been a method for using contextual definitions together with the rest of logic and set theory; "[t]o account for the external world as a logical construct of sense data..." (Quine 1969, p. 74, quoting Russell). This challenge was taken up by Rudolf Carnap, who said of Russell's influence (after reading Russell's *Our Knowledge of the External World*):

> I felt as if this appeal [for the application of logic to philosophical problems] had been directed to me personally. To work in this spirit would be my task from now on! And indeed henceforth the application of the new logical instrument for the purposes of analyzing scientific concepts and of clarifying philosophical problems has been the essential aim of my philosophical activity. (Carnap 1963, p. 12)

Quine sees Carnap's *Aufbau* as the most successful attempt, along Russellian lines, to complete the conceptual programme. Since Quine dedicates much space to its discussion, and bases his argument for the naturalisation of epistemology on the failure of Carnap to complete the project of the *Aufbau*, ~~so shall we~~ I shall also examine the project in detail.

C. The project of Carnap's *Aufbau*

Rudolf Carnap in his 1928 *Der Logische Aufbau der Welt*, following in the spirit of Hume and Russell, attempted to provide a means to translate "all sentences about the world into terms of sense data, or observation, plus logic and set theory." (Quine 1969, p. 74) Carnap did not, however, assume that the "stream of experience is composed of determinate, discrete elements," (Carnap 1928 [1967], p. 109). He therefore sought a means to divide up the stream of experience into parts such as sounds, smells and colours. He proposed a method of "quasi analysis" to divide the stream of experience into components by grouping them through a subjective similarity relation. The separate groups of experiences so formed were called by Carnap "similarity circles." The similarity circles provided the basis for the definition of constituents (or as Carnap called them, "quasi constituents") of the un-individuated experiences, in that experiences which are, say, visual were, Carnap claimed, similar to other visual experiences. (In fact, Carnap's procedure was more complex, since he defined the relation of similarity in terms of "recollection of similarity" since he thought this was "epistemically more fundamental" than other relations among experiences. (ibid, p. 127) I shall focus on similarity, since little turns on the difference between similarity and its recollection.)

Carnap gave as an example of quasi analysis the analysis of a musical chord. According to Carnap, when we hear a c-major chord we do not hear its constituent notes c, e, and g. But, the c-major chord may be said to be similar to all other chords which contain c, as it is similar to chords which also contain e, and which contain g, and so the c-major chord contains as quasi constituents c, e, and g: "Since the chord c-e-g is an element of similarity circles c, e, and g, we assign to it these three classes, namely, c, e, g, as quasi constituents." (Carnap 1928 [1967], p. 115)

It is evident that the similarity relation cannot be unique: c-major chords are similar in some respects to any other musical chord, to motions made on any instrument, to mathematical entities, and so on. Likewise, any coloured object is similar to other coloured objects, to objects of the same shape, to objects of the same mass, and so on. Hence, the similarity relation could not be used to create a category that contains only, say, acoustic or colour experiences. In fact, it seems to be generally the case that similarity relations cannot provide the basis for uniquely analysing experiences into certain categories, since everything is, in some respect, similar to anything else. If we want a unique similarity relationship, we must define it on subjective grounds: however, it is not clear that Carnap realised this. Since Quine's arguments against Carnap do not depend on Carnap's having assumed an objective similarity relationship, I shall interpret Carnap as having meant subjective similarity.

The aim of the *Aufbau* was to define the external world by means of definite descriptions referring to quasi constituents. Thus, all basic elements were to be ordered only by the relation of "recollection of similarity," and so all statements about the world were to be translated into a language of set theory, logic, and observation (hence the title, "The Logical Construction of the World"). Carnap only fully carried out his project up to the definition of colours, only indicating in outline how the rest of the project might be fulfilled.

For Quine, this is an impressive attempt to complete the conceptual programme. The advantages, according to Quine, of Carnap's *Aufbau*, had it been completed, would be first that it would advance the conceptual programme as far as possible. Since Quine holds that the reduction of mathematics to set theory and logic enhances overall clarity, despite the unclarity of set theoretic terms, because "of the interrelations that emerge," a parallel result might also be forthcoming from the Carnapian conceptual programme: "...such constructions would deepen our understanding of our discourse about the world, even apart from questions of evidence; it would make all cognitive discourse as clear as observation terms and logic and, I must regretfully add, set theory." (Quine 1969, p. 75)

The second advantage was to do with minimising the ontological foundations necessary for science, and so reducing the number of concepts in need of justification. The completed *Aufbau* would show that science could in principle be performed with the theoretical concepts of observation, set theory and logic alone, "...it would show all the rest of the concepts of science to be theoretically superfluous. It would legitimize them—to whatever degree the concepts of set theory, logic, and observation are themselves legitimate—by showing that everything done with the one apparatus could in principle be done with the other." According to Quine, "this would be a great epistemological achievement...", presumably because the foundations of science, would then be as firm as logic. (Quine 1969, p. 76) Moreover, the *Aufbau* would have been ontologically parsimonious, a great boon to those who, like Quine, "have a taste for desert landscapes." (Quine 1953 [1980], p. 4)

However, Quine denies that Carnap's programme, if successful, and in spite of reducing the number of entities in need of justification, could advance the doctrinal programme, since the translation of science into foundational terms would not show how to prove scientific statements from observation sentences: "...the mere fact that a sentence is couched in terms of observation, logic, and set theory does not mean that it can be proved from observation sentences by logic and set theory." (Quine 1969, p. 74) Quine, presumably following Hume, notes that a proof of the truth of any statement could never be given, since even "[t]he most modest of generalizations about observable traits will cover more cases than its utterer can have had occasion actually to observe." (Quine 1969, p. 74) According to Quine, we can only test our construction of the world against sense data—we will never be able to deduce logically our construction of the world from sense data.

The doctrinal programme, as defined by Quine, is concerned with "the hope of certainty," (Quine 1969, p. 74) but this, he says, is a forlorn hope:

> 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 its doctrinal side; but that quest was seen as a lost cause. To endow the 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 1969, p. 74, italics mine)

However, a modified doctrinal programme might be advanced by aiming for something less than certainty: "...such constructions could be expected to elicit and clarify the sensory evidence for science, even if the inferential steps between sensory evidence and scientific doctrine must fall short of certainty." (Quine 1969, p. 74-75) This is, perhaps, a reference to inductive probabilistic reasoning. As I shall show in section I, Quine was aware of Carnap's programme to define probability as degree of entailment, and I shall show that Quine himself views probabilities as a means of representing uncertainty. Quine's claim that we might discover "inferential steps" which "fall short of certainty," might refer to an attempt to partially infer the laws of science from the evidence of our senses (as explicated by the Carnapian programme).

To sum up, the history of the doctrinal and conceptual programmes from Descartes to Carnap show, first, that the doctrinal programme cannot achieve certainty (Hume demonstrated this), and second that the conceptual programme was most strongly advanced by Bentham, Russell and Carnap, through the use of definite descriptions. Quine seems to believe that this history contains the most successful empiricist attempts to provide foundations for epistemology. However, according to Quine, all attempts to complete the doctrinal programme have failed. It is also clear that Carnap's conceptual programme likewise failed. Quine argues that no similar programmes will be successful: I examine Quine's reasons for holding the former in the next section.

D. The failure of the conceptual programme (the indeterminacy of translation)

As Quine notes, Carnap himself conceded that the project of the *Aufbau* was uncompletable. In fact, there are two reductions which fail, and I shall examine these in turn. Quine argues, on the basis of his indeterminacy of translation thesis that any attempt to reduce scientific language to an observation language is bound to fail. I will examine this argument in section D.1, after discussing the actual failures of Carnap's constructions.

In his 1931 article "Testability and Meaning", Carnap considered whether it was possible to define "disposition-concepts," such as 'visible' or 'soluble.' Taking 'soluble' as an example, and three predicates, $Q_3(x)$, "x is soluble in water", $Q_1(x,t)$, "the body x is placed in water at time t", and $Q_2(x,t)$, "the body x dissolves at time t", Carnap asked whether "x is soluble in water" could be defined as

$$D: Q_3(x) \leftrightarrow \forall(t)[Q_1(x,t) \rightarrow Q_2(x,t)].$$

His answer was that it could not: consider an object, c, which is not soluble in water, which has never been put in water, and which has subsequently been completely destroyed, say, a match which has been completely burnt. $Q_3(c)$ is false, because wood is not soluble in water. But so is $Q_1(c,t)$, for any value of t, since the match has never been placed in water, nor will it ever be placed in water, which makes the right-hand side of the definition true. Therefore, $Q_3(c)$ must be true, and the match is soluble. According to Carnap, this shows that disposition concepts can never be defined: "'Q_3' cannot be defined by D, nor by any other definition." (Carnap 1936, p. 440)

Instead of reducing disposition statements to statements only about basic empirical concepts by means of definitions, Carnap proposed that we use "reduction sentences" to interpret disposition concepts. Reduction sentences are of the form

$$\forall(x)\forall(t)[Q_1(x,t) \rightarrow (Q_3(x) \leftrightarrow Q_2(x,t))],$$

and are meant to avoid the problem above. While they will not in general show us how to eliminate disposition concepts from our language entirely, they show how disposition concepts are related to non-disposition concepts. Carnap noted that "terms introduced in this way have the disadvantage that in general it is not possible to eliminate them, i.e., to translate a sentence containing such a term into a sentence containing previous terms only." (Carnap 1936, p. 443) Quine points out the importance of the difference between implications (which are given by reduction sentences) and equivalences (given by definitions): If we have equivalences of the type given in the *Aufbau*, on which the left hand side is the scientific statement to be reduced, and the right hand side contains only observational terms, we can eliminate, if we wish, all scientific language in favour of talk about observation terms, sets, and logic. We cannot, in general, do this with reduction sentences, as Carnap noted, and so the programme of translating all of science into statements about observables cannot succeed, and so cannot demonstrate that theoretical terms are eliminable in favour of observation terms. Of theoretical terms, we still require, at least, dispositional terms.

Another area in which Carnap failed to reduce scientific concepts to sense data is to be found in his method of assigning of colours, or any such "quality" to particular space-time points, that is, of reducing the statement "quality x is at space-time point y" to a statement only about sense data and sets. Carnap provided only a sketch, in terms of a list of twelve desiderata (what Quine calls "canons"), of how the assignment of qualities to space-time points was to proceed. These assumptions are in fact quite substantial, presupposing a fairly large knowledge of physics. For example, that "The speed of light, c, is constant and very large. Thus, the light rays are very nearly the straight lines of a momentary space", or "We shall assume, if there is no reason to the contrary, that each point of the outside world retains at the other times the same or as similar as possible a color as that with which it was seen at one time." And another of Carnap's desiderata, which I will refer to again, is that we were initially to assume that in our observations, "the optical medium between the eye and the seen things can generally be considered homogeneous" (Carnap 1928 [1961], §127).

Carnap admits that "[a]ctually, without clearly realizing it, I already went beyond the limits of explicit definitions in the construction of the physical world. For example, for the correlation of colors with space-time points, only general principles, but no clear operating rules were given (§127)." (Carnap, introduction to the second edition of the *Aufbau*, p. viii). Quine follows Carnap in identifying this as an insurmountable barrier for the programme of describing the physical world solely in terms of sense data and definite descriptions, although his reasons are not clear: "The crucial point comes where Carnap is explaining how to assign sense qualities to positions in physical space and time. These assignments are to be made in such a way as to fulfil, as well as possible, certain desiderata which he states and with growth of experience the assignments are to be revised to suit. This plan, however illuminating, does not offer any key to translating the sentences of science into terms of observation, logic, and set theory." (Quine 1969, p. 77). Quine makes the same point in his "Two Dogmas of Empiricism": "Carnap did not seem to recognize, however, that his treatment of physical objects fell short of reduction not merely through sketchiness, but in principle." (Quine 1953 [1980], p. 40) Quine remarks that Carnap provides no means by which "a statement of the form 'Quality q is at point-instant x;y;z;t' could ever be translated into [his] initial language of sense data and logic. The connective 'is at' remains an added undefined connective; the canons counsel in its use but not in its elimination." (Quine 1953 [1980], p. 40)

The problem with adequately defining the notion of a quality being at some space-time point seems to be that it requires scientific information as to the actual relation between the view and what is viewed. If our visual image represents some contemporary and local physical state, we need to assume that nothing is intervening to deceptively shift the subjective place of the object in relation to its real place. For example, sticks that at first look and feel straight, when placed in a glass of water appear bent. This is because we may not always assume "the optical medium between the eye and the seen things" is "homogeneous": we require information to determine when this assumption may be made and when it may not be made. However, if such information is required to define "is at", then "is at" is not reduced, but is defined in terms of the science it is intended to reduce.

Hence, for the two reasons I have cited, which Carnap himself conceded, Carnap's conceptual programme cannot succeed. The conceptual programme was an attempt to clarify scientific discourse by translating all sentences of science into sentences consisting only of logical and set theoretic terms in the form of definite descriptions of objects referring only to phenomenal experience. Since disposition concepts are ineliminable, and an essential part of science, we cannot translate scientific statements into a pure observation language, and hence cannot show that science relies only on logic, set theory, and observation.

Quine's criticism goes even further than pointing out that Carnap himself failed; he argues on the basis of his famous thesis of the indeterminacy of translation that no translational reduction is possible. The argument is important because, if correct, it would first show that the traditional conceptual programme was misconceived, and second that the doctrinal programme would also not be possible, since its role is to justify the epistemic relations between statements revealed by the conceptual programme. Hence I shall examine Quine's argument in detail.

1. The indeterminacy of translation

Quine argues that Carnap's programme fails because of the alleged indeterminacy of translation, the thesis that there is no unique way of translating one language into another, and that no translation is more justified than any other. In this section I examine the grounds for Quine's thesis and why Quine thinks it undermines the conceptual programme. I argue that Quine's arguments are unsound, and that even if the indeterminacy of translation were a fact, it would not undermine the Carnapian conceptual programme. This is an important point, because, as I shall show, Quine argues from the impossibility of completing the conceptual programme to the necessity of adopting what he calls naturalised epistemology.

According to Quine, the indeterminacy of translation follows from a verification theory of meaning and the Duhem thesis. Quine traces the verification theory of meaning to Peirce, who, Quine claims, held that "the very meaning of a statement consists in the difference its truth would make to possible experience." (Quine 1969, p. 78) Quine also phrases the verification theory of meaning as the doctrine that "the meaning of a sentence turns purely on what would count as evidence for its truth..." (Quine 1969, p. 80) The

Duhem thesis may be illustrated thus: assume that we wish to test a part of chemical theory by checking its predictions. The theory by itself will, typically, yield no empirical predictions. For example, it may seem that the theory predicts that a certain chemical reaction will yield a certain amount of heat, and that we may test the theory by performing the chemical reaction and measuring the heat released. However, the prediction only follows from the theory conjoined with auxiliary assumptions having to do with, for example, the cleanliness and calibration of the equipment, the purity of the chemicals involved, and so on. Thus, the chemical theory, and scientific theories in general, are not testable without auxiliary assumptions. This thesis is usually held to be a problem for a falsificationist methodology, since it highlights the fact that there are theories that are generally considered to be scientific but which are not falsifiable; Quine also holds that the Duhem thesis applies to the confirmation of theories, that is, only theories in conjunction with auxiliary assumptions are verifiable or falsifiable:

> The dogma of reductionism survives in the supposition that each statement, taken in isolation from its fellows, can admit of confirmation or infirmation at all. My countersuggestion, issuing essentially from Carnap's doctrine of the physical world in the *Aufbau*, is that our statements about the external world face the tribunal of sense experience not individually but only as a corporate body. (Quine 1953 [1980], p. 41)
>
> Suppose an experiment has yielded a result contrary to a theory currently held in some natural science. The theory comprises a whole bundle of conjoint hypotheses, or is resoluble into such a bundle. The most that the experiment shows is that at least one of those hypotheses is false; it does not show which. It is only the theory as a whole, and not any one of the hypotheses, that admits of evidence or counter-evidence in observation and experiment. (Quine 1970a, p. 5)

Quine uses the Duhem thesis, by conjoining it with an empiricist criterion of meaning, to support his thesis of *meaning holism*, that is, that only "sufficiently inclusive" parts of language have meaning. Since practically all scientific theories require auxiliary assumptions in order to imply confirmable or falsifiable consequences, any empirical consequences of a theory will depend on these assumptions. If the meaning of a theory is its empirical consequences, and a theory only has empirical consequences in conjunction with auxiliary assumptions, it follows that such a theory by itself has no meaning:

> ...the typical statement about bodies has no fund of experiential implications it can call its own. A substantial mass of theory, taken together, will commonly have experiential implications; this is how we make verifiable predictions....
>
> Sometimes also an experience implied by a theory fails to come off; and then, ideally, we declare the theory false. But the failure falsifies only a block of theory as a whole, a conjunction of many statements. The failure shows that one or more of those statements is false, but it does not show which. The predicted

> experiences, true and false, are not implied by any one of the component statements of the theory rather than another. *The component statements simply do not have empirical meanings, by Peirce's standard [the verification theory of meaning]; but a sufficiently inclusive portion of theory does.* (Quine 1969, p. 79)

Thus, according to Quine, only large parts of scientific theory ("sufficiently inclusive portion[s]") have meaning. For Quine, meaning holism only holds for theoretical, not observation sentences, as we shall later see. Quine's views on just how large these parts are have changed somewhat from the comprehensive holism of "Two Dogmas of Empiricism", in which Quine held that "The unit of empirical significance is the whole of science." (Quine 1953 [1980], p. 42) In his 1980 foreword to *From a Logical Point of View*, Quine seems to endorse a more modest meaning holism by stating that in practice portions of science, as opposed to the whole, have empirical content:

> The holism in "Two dogmas" has put many readers off, but I think its fault is one of emphasis. All we really need in the way of holism, for the purposes to which it is put in that essay, is to appreciate that empirical content is shared by the statements of science in clusters and cannot for the most part be sorted out among them. Practically the relevant cluster is indeed never the whole of science... (Quine [1980], p. viii)

However, this modest meaning holism applies only "practically." Quine claims that "legalistically," radical holism is defensible, since mathematics and logic could be revised instead of a theory, as has been seen in the development of quantum and intuitionistic logics (Quine 1970a, p. 100, 1953 [1980], p. 43), and mathematics and logic are shared by all branches of science:

> ...this structure of interconnected sentences is a single connected fabric including all sciences, and indeed everything we ever say about the world; for the logical truths at least, and no doubt many more commonplace sentences too, are germane to all topics and thus provide connections. However, some middle-sized scrap of theory usually will embody all the connections that are likely to affect our adjudication of a given sentence. (Quine 1960, pp. 12-13)
>
> And how wide is a theory? No part of science is quite isolated from the rest. Parts as disparate as you please may be expected to share laws of logic and arithmetic, anyway, and to share various common-sense generalities about bodies in motion. Legalistically, one could claim that evidence counts always for or against the total system, however loose-knit, of science. Evidence against the system is not evidence against any one sentence rather than another, but can be acted on rather by any of various adjustments. (Quine 1970a, p. 5)

(Quine makes the same point in his 1986, pp. 619-620).

The thesis of the indeterminacy of translation is the application of meaning holism to translation. Since translations must preserve meanings, and since meaning, according to Quine, is a property of 'large' statements, a translation between short sentences in different languages is not possible: only "sufficiently inclusive" parts of theories may be translated. But according to Quine, this translation of blocks introduces an indeterminacy because there will be many different ways of translating the blocks of theory by differently interpreting the sentences which compose the theories. These theories may differ radically in terms of, say, basic entities, and yet still be compatible with the same empirical evidence:

> ...it is to be expected that many different ways of translating the component sentences, essentially different individually, would deliver the same empirical implications for the theory as a whole; deviations in the translation of one component sentence could be compensated for in the translation of another component sentence. Insofar, there can be no ground for saying which of two glaringly unlike translations of individual sentences is right. (Quine 1969, p. 80)
>
> If we recognize with Peirce that the meaning of a sentence turns purely on what would count as evidence for its truth, and if we recognize with Duhem that theoretical sentences have their evidence not as single sentences but only as larger blocks of theory, then the indeterminacy of translation of theoretical sentences is the natural conclusion." (Quine 1969, pp. 80-81)

Quine acknowledges that despite the indeterminacy of translation, it is possible to translate scientific language into observational language, albeit a translation of large parts of theories as opposed to individual hypotheses: "If we can aspire to a sort of *logischer Aufbau* at all, it must be to one in which the texts slated for translation into observational and logico-mathematical terms are mostly broad theories taken as wholes, rather than just terms or short sentences. The translation of a theory would be a ponderous axiomatization of all the experiential difference that the truth of the theory would make." (Quine 1969, p. 79) Such a translation would certainly not be the type of translation we ordinarily encounter, and while it is clear that Quine believes that such a holistic translation is not what Carnap would have wanted, it is not clear why he feels that the indeterminacy of translation is fatal to the Carnapian programme. As noted earlier, in section C, Quine envisaged three benefits resulting from the completion of the *Aufbau*: that it would clarify science; that it would make science ontologically parsimonious; and that it would advance a modified doctrinal programme. I shall now argue that these benefits are not lost due to the indeterminacy of translation.

It might seem that a complete *Aufbau* would not clarify science if, as Quine claims, there can be no unique translation of a theory into observation terms. The Duhem thesis, taken together with the verification theory of meaning, implies the meaningless of individual statements. A further argument is needed to demonstrate that the meaninglessness of individual statements implies that there can be no unique translations. Quine's argument for this is that there would be many ways of matching up the component sentences of two theories so that they have the same empirical consequences. Hence terms could be translated differently than they usually are, and still the empirical consequences of the theory would be preserved. According to Quine, this precludes term-by-term translation. Quine makes the seemingly critical point about holistic translation that it is "queer" and seems not to be translation at all: "We might better speak in such a case not of translation but simply of observational evidence for theories; and we may, following Peirce, still fairly call this the empirical meaning of the theories." (Quine 1969, pp. 79-80) However, Quine gives no arguments for his claim that the indeterminacy of translation rules out term-by-term translation. Consider the case in which we have translated a block of theory into a large sentence or set of sentences in the observation language. Nothing in Quine's argument precludes us then, perhaps arbitrarily, correlating individual 'small' sentences in the blocks with individual 'small' sentences in the observation language. The only stricture on such a linking would be that it preserve the empirical content of the whole scientific theory. Quine's argument, even if successful, would merely show that there is no unique or privileged way of linking up the sentences of the theories. However, Quine seems to have acknowledged that the impossibility of a privileged translation poses no difficulty for Carnap's programme: Quine tells us that Carnap sought a rational reconstruction, and that "[a]ny construction of physicalistic discourse in terms of sense experience, logic, and set theory would have been seen as satisfactory if it had made the physicalistic discourse come out right." And, further, Quine acknowledged that there would be no unique construction: "If there is one way, there are many..." Nevertheless, Quine tells us that "any would be a great achievement." (Quine 1969, p. 75) Thus, even granted the indeterminacy of translation, the first advantage, that of translating scientific language into a clearer language, does not seem to be lost.

The second advantage, ontological parsimony, would likewise seem to be retained. Granted, there would be no unique, parsimonious translation of science, but, as just noted, this was not Carnap's aim. The purpose of the construction was to show that the theoretical entities of science could in principle be eliminated in favour of logic, sets, and observation terms. Whether the indeterminacy of translation thesis is true or not, the goal of translating theories into an observation language in a parsimonious manner is surely unaffected. Hence science could, in principle, be carried out solely within an observation language, that is, that science could be done in an ontologically parsimonious fashion.

I shall now argue that the third projected advantage of the *Aufbau*, that of making clear what inferential steps, short of certainty, are needed to establish knowledge of the outside world, is also not compromised by the indeterminacy of translation. To argue against Quine on this point, I shall take issue with his confirmational holism, and thus with the thesis of the indeterminacy of translation. I therefore devote a section to this argument.

2. Against confirmational holism

It is worth quoting Quine again as to the nature of the third benefit to be conferred by a completed *Aufbau*: "...such constructions could be expected to elicit and clarify the sensory evidence for science, even if the inferential steps between sensory evidence and scientific doctrine must fall short of certainty." (Quine 1969, p. 74-75) If empirical support alone determines confirmation, which certainly seems to be Quine's view, then the fact that individual statements have no empirical implications might seem to imply that they cannot be confirmed. We could not specify any "inferential steps" because there would be no "steps", no series of statements, but only one large statement, an entire theory. (If this were the case, it would be a consequence of the Duhem thesis, and not of the indeterminacy of translation.) However, the argument is dependent on a confirmation theory. I shall now argue that Quine's claim does not hold for all confirmation theories.

Consdier a Bayesian point of view, and suppose, for simplicity's sake, that we have a theory that can be naturally represented as $T \wedge A$, where T is, say, the above-mention theory of chemical reactions, and A a set of auxiliary assumptions concerning the competence of the laboratory assistants, etc. Further suppose that neither T nor A is 'large' enough itself to have empirical consequences, but that when conjoined they imply some observation E. Quine's claim is then that E cannot confirm T, which in Bayesian terms means that $p(T|E) - p(T) = 0$, or, equivalently, that $p(T|E) = p(T)$ (that is, T and E are independent). This, however, will not in general be the case. By the probability calculus,

$$p(T|E) = p(T \wedge A|E) + p(T \wedge \neg A|E).$$

By hypothesis, $p(T \wedge A|E) > p(T \wedge A)$. It is plausible to assume that the theory and the assumptions could be such that, while, $p(T \wedge A|E)$ is very high, $p(T \wedge \neg A|E)$ is very low. This need merely be a feature of a well-designed experiment. There then are probability distributions such that E does in fact confirm T. Since T would be separately confirmable, it would have separable meaning. Hence a Bayesian account of confirmation does not support meaning holism.

In fact, as I shall now argue, Quine would agree with the assumptions I have made in this argument. Quine acknowledges that in practice scientists do assign different epistemic weights to different hypotheses: "...the separate sentences of science and common sense do in practice seem after all to carry their separate empirical meanings." (Since Quine holds that the meaning of a sentence is determined by its empirical evidence, separate empirical meaning is the same as separate epistemic weight.) However, according to Quine the assigning different epistemic weights by scientists is "misleading, and explicable." (Quine 1970, p. 7) He considers the case in which a molecular biologist predicts some occurrence, which does not happen. According to Quine, the scientist "is apt to scrutinize for possible revision only the half dozen beliefs that belonged to molecular biology rather than tamper with the more general half dozen having to do with logic and arithmetic and the gross behavior of bodies." (ibid, p. 7) According to Quine the scientist employs "a reasonable strategy—a maxim of minimum mutilation." (ibid) Yet Quine asserts that the "reasonable" practice of considering a hypotheses "as more tentative and suspect than other parts of the theory", which is a result of "connections of varying strengths that incline us to affirm or deny some sentences when affirming or denying others", "does not show how to allocate separate empirical evidence to separate sentences." (ibid) If the Bayesian is correct, however, Quine is wrong, and the tendency of scientists to believe some sentences more firmly than others does show us how to distribute empirical support over those sentences.

Quine's holistic view of meaning also has a strange consequence, namely, that meaningless statements can have truth values. In the above example, $T \wedge A$ is meaningful. It surely could also be true. But $T \wedge A$ entails T. Hence T could be true, but meaningless. Quine also characterises the scientist's suspicion of a single hypothesis as justifiable. Hence, for Quine, it is justifiable to believe some meaningless statements more firmly than other meaningless statements. This peculiar result seems to be another argument against Quine's holistic verification theory of meaning.

It might be objected that I have contradicted myself in this section: in Chapter 3 I argued that the Dutch Book and the Money Pump arguments are invalid, but in the argument above I employ Bayesian confirmation theory. However, the argument in Chapter 3 is not that Bayesian confirmation theory is wrong, but that it is not justified apriori. Also, I shall show in section I.1 that Quine accepts a Bayesian account of confirmation.

I conclude that the indeterminacy of translation does not undermine the *Aufbau* project. Even if Quine's thesis is true, it does not show the impossibility of an ontologically parsimonious translation of science into an observation language. Further, it seems to me that the thesis of meaning holism, on which the indeterminacy of translation claim is based, is untenable, and has the strange consequence that meaningless statements may be believed, and indeed, even have truth values.

Perhaps another argument would serve Quine as well: it is surely impossible to effect a Carnapian translation. Consider disposition terms—it may be possible, by creating very complex logical constructions, to explain them in an extensional observation language. This seems, however, very unlikely. Disposition statements seem inherently intensional and irreducible to extensional statements, and so the project of the *Aufbau* seems doomed to failure. This does not, however, serve as an in-principle argument, which is presumably what Quine wants. Quine's goal was to demonstrate that the traditional epistemological programme, as he conceives it, has failed, and must always, fail. Once this is established, Quine claims, then psychology can provide a substitute for the traditional epistemological programme.

E. (Part of) Epistemology Naturalised

In this section I argue that Quine aims to replace only the conceptual programme with a psychological account of learning. (His account of the doctrinal programme is much more complex, and I shall discuss it in the following sections.) According to Quine, the Carnapian programme has failed, but its aim of clarifying the notion of observation, can still be pursued. However, for Quine, an aposteriori account is required.

Quine gives two reasons why one might prefer a Carnapian rational reconstruction of the relation of sense data to scientific theories over a psychological account. First, "We should like to be able to translate science into logic and observation terms and set theory..." Psychology, obviously, cannot provide such a translation: "If psychology itself could deliver a truly translational reduction of this kind, we should welcome it; but certainly it cannot, for certainly we did not grow up learning definitions of physicalistic language in terms of a prior language of set theory, logic, and observation" (Quine 1969, p. 76). Quine refers to translational reduction as "the last remaining advantage that we supposed rational reconstruction to have over straight psychology." (Quine 1969, p. 78). But, as Quine has argued, the Carnapian programme has failed. Hence, psychological investigations provide an alternative for the programme of clarifying how theories and evidence are related: "If all we hope for is a reconstruction that links science to experience in explicit ways short of translation then it would seem more sensible to settle for psychology. Better to discover how science is in fact developed and learned than to fabricate a fictitious structure to a similar effect." (Quine 1969, p. 78) Quine's claim that the Carnapian conceptual programme was a failure is uncontroversial, and seems to me correct, although I dispute Quine's claim that the indeterminacy of translation is the reason for this failure. The claim that scientific terms could be reduced to other empirical terms by employing the social sciences seems justified, and if one is motivated to "clarify" scientific terms by defining them in empirical terms, and if these terms cannot be defined in a language consisting solely of logic and observation, then an obvious alternative is to define them in a language which is supplemented by terms from, say, psychology. This suggestion, however, only applies to the conceptual programme.

Quine's second reason for preferring a Carnapian programme to a psychological one is that adopting the latter seems to be reasoning in a circle:

> But why all this creative reconstruction, all this make-believe? The stimulation of his sensory receptors is all the evidence anybody has had 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? Such a surrender of the epistemological burden to psychology is a move that was disallowed in earlier times as circular reasoning. If the epistemologist's goal is validation of the grounds of empirical science, he defeats his purpose by using psychology or other empirical science in the validation. (Quine 1969, pp. 75-76)

I argued above that it would indeed be circular reasoning to adopt an empirical approach to methodology. Quine seems to acknowledge this too, and seems to wish to replace only the conceptual programme with psychology, not the doctrinal programme, which as I argued above he seems to view as hopeless. Quine argues that if we are no longer "dreaming of deducing science from observations," then we need not be concerned by the charge of circular reasoning:

> However, 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 1969, p. 76)

> This interplay [between natural science and epistemology] is reminiscent again of the old threat of circularity, but it is all right now that we have stopped dreaming of deducing science from sense data. (Quine 1969, p. 83)

The phrase "the link between observation and science" is, like the phrase "how evidence relates to theory", ambiguous as to whether the link is doctrinal or conceptual, and a further ambiguity is introduced by Quine's claim that his own argument is non-circular because epistemologists have abandoned the programme of "deducing science from sense data." Two interpretations suggest themselves: first, that the use of psychology is not circular since it is intended only for the conceptual, not the doctrinal programme; second, that psychological considerations are important in a doctrinal programme that does not seek to provide deductive justifications of theory: this second interpretation leaves open the possibility of non-deductive justificatory strategies.

The authors cited at the beginning of this chapter seem to opt for the first interpretation. I shall argue that Quine offers a more complicated position: the conceptual programme should be a purely empirical enterprise, and the doctrinal programme a conventional enterprise but with an empirical element. However, this interpretation requires a close examination of Quine's views on naturalised epistemology. Therefore, in section F, I present Quine's account of the naturalised conceptual programme. I will demonstrate that this programme relies on a particular confirmation theory. I shall then examine this confirmation theory in section G, and argue in section H that Quine accepts the confirmation theory on conventionalist grounds.

My reading of Quine may seem implausible at first: why, it might reasonably be asked, did Quine title his main work on the subject "Epistemology Naturalised" if he meant "Part of Epistemology Naturalised"? I suggest that this is because Quine is more interested in the conceptual programme. I argued above that Quine views the doctrinal programme as having failed: he also seems to believe that this programme, as an apriori discipline, cannot succeed. For example, under the entry "Knowledge" in his 1987 *Quiddities*, Quine defines knowledge as "justified true belief." Quine notes the well-known difficulties with this definition such as the Gettier problem. This problem turns on the fact that even given a very good justification for a belief, if the belief turns out to be true by accident, the justified true belief does not seem to count as knowledge. For example, suppose I have participated in a quiz show for a week. I have won quite easily each day, and on the last day of the week am facing an opponent I vanquished with particular ease earlier. Suppose further that I have three coins in my pocket. I certainly have good reason for believing that the winner of today's quiz will have exactly three coins in his pocket. In fact, it turns out I lose, but it so happens that my opponent also has exactly three coins in his or her pocket. Therefore my belief was both justified and true—but it is not knowledge as ordinarily understood. From this, Quine concludes that knowledge is a "bad job":

> I think that for scientific or philosophical purposes the best we can do is give up the notion of knowledge as a bad job and make do rather with its separate ingredients. We can still speak of a belief as true, and of one belief as firmer or more certain, to the believer's mind, than another. There is also the element of justification, but we saw its limitations. (Quine 1987, p. 109)

If the notion of knowledge is a bad job, then so is its study. However, Quine has suggested that we study the components of knowledge, and while this is a departure from "traditional epistemology," as Quine conceives it, it still aims for the clarification of epistemic concepts. It is therefore within the tradition of epistemology, and hence "epistemology, or something like it" (Quine 1969, p. 82) is naturalised. Quine does in fact go into detail about the naturalised conceptual programme, and it is to this I now turn.

F. An example of the conceptual programme naturalised: the observation sentence.

In this section I shall explicate Quine's notion of the observation sentence, placing it in the context of the conceptual programme. I focus on the observation sentence because, as I shall show, Quine considers it basic to both the conceptual and doctrinal programmes. He takes the notion of an observation sentence to be the basic notion of the naturalised conceptual programme. However, I shall argue, that Quine's account of the observation sentence requires a theory of confirmation. I shall also argue that Quine supplies the outline of a confirmation theory. In following sections I shall argue that this confirmation theory is properly placed within the doctrinal programme.

Quine's theory of the observation sentence is developed from his account of how a language is acquired by a translator and by an infant. I shall examine Quine's account of the former, that is, how a translator would proceed when faced with a language for which no interpreters (or chain of interpreters) are available. Since the linguist must start from scratch, Quine denotes this "translation of the language of a hitherto untouched people" as "radical" translation, (Quine 1960, p. 28)

1. The method of radical translation

The first step, Quine claims, in radical translation is to determine the native speakers' words for assent and dissent. According to Quine, this is done by correlating situations with native speakers' utterances. His example is the word "Gavagai," which, he supposes, is spoken, at least once, by a native in the presence of a rabbit. The translator notes the coincidence of the utterance of the word and the presence of the rabbit, and tests the conjecture that the word means "rabbit." When a rabbit happens by, the linguist says "Gavagai", noting the response of the native speaker. Similarly, the translator says "Gavagai" when no rabbits are available, and notes the words used in response. The linguist then conjectures that these responses, or some subset of them, determine assent or dissent to questions, that is, they correspond to "yes" and "no."

Quine acknowledges that this may be very difficult, since the significance of facial expressions and gestures, for example, change from culture to culture. (Quine 1960, pp. 29-30) Quine optimistically suggests that the translator can safely assume that he is almost always correct, and so the word which he hears in answer to his question means "yes", as opposed to "no" or even, "fool!" Quine also suggests that words evoking a "more serene" response mean "yes". Having developed a hypothesis about the words for "yes" and "no", the translator can then attempt to associate words with situations. For example, he would translate words like "gavagai" into "rabbit" if, whenever a rabbit is present and he says "gavagai" the native speaker responds "yes".

Quine's theory of translation is meant to serve as the basis for a purely empirical theory of meaning. There are two steps in the construction of such an account. The first is to relate native responses solely to stimuli:

> It is important to think of what prompts the native's assent to 'Gavagai?' as stimulations and not rabbits. Stimulation can remain the same though the rabbit be supplanted by a counterfeit. Conversely, stimulation can vary in its power to

> prompt assent to 'Gavagai' because of variations in angle, lighting, and color contrast, though the rabbit remain the same. In experimentally equating the uses of 'Gavagai' and 'Rabbit' it is stimulations that must be made to match, not animals. (Quine 1960, p. 31)

The second step is to equate the stimuli which prompt the utterance of a word with the meaning of the word:

> Certain sentences of the type of 'Gavagai' are the sentences with which our jungle linguist must begin, and for these we now have before us the makings of a crude concept of empirical meaning. For meaning, supposedly, is what a sentence shares with its translation; and translation at the present stage turns solely on correlations with non-verbal stimulation. (Quine 1960, p. 31)

(Quine takes 'Gavagai' to be a one word sentence, perhaps short for "Lo, a rabbit.") Quine calls this type of meaning "stimulus meaning," which he defines as the combination of "affirmative stimulus meaning", the stimuli which would prompt an affirmative response to a statement by a native speaker, and "negative stimulus meaning", the stimuli which would prompt a negative response. (Quine 1960, pp. 32-33)

According to Quine, his account of meaning supports the indeterminacy of translation thesis. For example, we may translate the native word "Gavagai" as "time-slice of a rabbit" or "undetached part of a rabbit": "Who knows but what the objects to which this term applies are not rabbits after all, but mere stages, or brief temporal segments, of rabbits? In either event the stimulations that prompt assent to 'Gavagai' would be the same as for 'Rabbit.'" (Quine 1960, p. 51-52)

An apparent difficulty with Quine's approach is that when the translator inquires as to the meaning of a specific word, he cannot be sure that the native has understood that he has been asked a question, nor that the native's nod of the head, say, is an affirmative answer. As Quine correctly recognises, a translator's conclusions about the meanings of words, for example words used for dissent or assent, are "inconclusive" and a "working hypothesis," which may, in certain circumstances, have to be abandoned: "If extraordinary difficulties attend all his subsequent steps, the linguist may decide to discard that hypothesis [i.e., the hypothesis concerning which native responses correspond to "yes" and "no"] and guess again." (Quine 1960, pp. 29-30) It seems clear that the construction of a translation manual is a fallible enterprise, even at the most basic level of determining native assent or dissent.

One reason for the fallibility of any translation is that we can never pin down with complete certainty which stimuli prompt assent and dissent: "In taking the visual stimulations as irradiation patterns we invest them with a fineness of detail beyond anything that our linguist can be called upon to check for." (Quine 1960, p. 31) Quine claims that the translator can "reasonably conjecture" that these stimuli are like the stimuli which would cause the translator himself to assent or dissent: "But this is all right, he can reasonably conjecture that the native would be prompted to assent to 'Gavagai' by the microscopically same irradiations that would prompt him, the linguist, to assent to 'Rabbit', even though this conjecture rests wholly on samples where the irradiations concerned can at best be hazarded merely to be pretty much alike." (Quine 1960, p. 31) And, even after the signs of assent and dissent have been reasonably well established, the translation of words such as "Gavagai" will similarly be conjectural. Quine suggests that we use what we know of our own stimulus meanings to judge the native's:

> Let us suppose the linguist has settled on what to treat as native signs of assent and dissent. He is thereupon in a position to accumulate inductive evidence for translating 'Gavagai' as the sentence 'Rabbit.' The general law for which he is assembling instances is roughly that the native will assent to 'Gavagai?' under just those stimulations under which we, if asked, would assent to 'Rabbit?'; and correspondingly for dissent. (Quine 1960, p. 30)

This hypothesis is, of course, not certain: Quine says that the construction of a translation manual is a matter of "accumulat[ing] inductive evidence." This is a crucial point. Quine aims to use his account of radical translation as the basis for a naturalised epistemology. But this account rests on the notion of evidence, and on induction. I shall show in the next section that Quine does have theory of evidence, based on what he calls "observation sentences," and that these sentences are, according to Quine, learned by induction.

2. The observation sentences

An observation sentence is defined by Quine as "one on which all speakers of the language give the same verdict when given the same concurrent stimulation." (Quine 1969, pp. 86-87. Quine also calls them the Protokollsätze, following Carnap and Neurath (Quine 1969, p. 85) Observation sentences are, for Quine, a subclass of what he calls "occasion sentences," which are sentences assented to or dissented from depending on present stimuli. As Quine puts it, occasion sentences are those "that admit of verdicts and truth values not once for all but from one occasion of utterance to another, depending on what is going on in the neighborhood." (1977 [1979], p. 156) An example of an occasion sentence is "There is a rabbit here": if I say "There is a rabbit here" someone would say "Yes" or "No" depending on whether they saw something that yields the same stimuli as a rabbit nearby. Other examples given by Quine are 'It's raining', 'This is red', and 'He owes me money'. However, "historical truths... scientific hypotheses,... credos,... slogans" are excluded, presumably because they are not assented to or dissented from on the basis of presented stimuli. (Quine 1977 [1979], p. 156) These are instead examples of "standing sentences", such as "Sugar is sweet," which are assented to or dissented from regardless of present stimulation.

Occasion sentences may be ranked according to the proportion of native speakers assenting to them in the presence of a specific stimulus. Occasion sentences which prompt universal assent or dissent from competent speakers subjected to the same stimuli are what Quine calls observation sentences. (Quine's definition of an observation sentence has the counter-intuitive consequence of classifying certain false sentences as observational: for example, in the Middle Ages, "There is a witch" would meet with near universal assent, if said near to a sufficiently strange looking person. I shall not purse this criticism, however.) Quine's definition of an observation sentence and his account of translation imply that observation sentences are the easiest to translate:

> [The observation sentence's] relation to meaning is fundamental,... since observation sentences are the ones we are in a position to learn to understand first, both as children and as field linguists. For observation sentences are precisely the ones that we can correlate with observable circumstances of the occasion of utterance or assent, independently of variations in the past histories of individual informants. They afford the only entry to a language. (Quine 1969, pp. 88-89)

Quine identifies ease of translation as a "mark of clarity," although a fallible one. In so doing, he claims to provide a naturalistic basis for the conceptual programme, clarity being defined in terms of breadth of agreement on the (stimulus) meaning of a statement. Thus, observation sentences are the clearest statements in a language. Since Quine defines meaning solely in terms of stimulation, universal assent or dissent as to the truth of a statement given the same stimuli is equivalent to universal agreement on the meaning of that statement. Hence such an agreed upon statement is maximally clear:

> The vehicle of communication is the sentence, and one mark of clarity of communication is agreement as to the truth of the sentence. This is a very fallible criterion, but it is a beginning.... The occasion sentences that pass this clarity test with high marks are what I call observation sentences. (Quine 1977 [1979], p. 156)

The criterion of universal assent leads naturally to a criterion of observability: for a sentence to count as an observation sentence, "all speakers of the language" must agree on its truth value. It also leads to a graded notion of observationality:

> ...an occasion sentence may be said to be the more observational the more nearly its stimulus meanings for different speakers tend to coincide.... Viewing the graded notion of observationality as the primary one, we may still speak of sentences simply as observation sentences when they are high in observationality. (Quine 1960, pp. 43-44)

As the last quote but one makes clear, sentences which command universal assent are highest in observationality, and so are observation sentences proper. It is, of coure, too much to demand that observation sentences command universal assent, and thus Quine defines an observation sentence as one to which *"all* members of the community, *nearly enough*, will say 'Yes' to it under the same stimulations, and *all* will say 'No' to it under the same stimulations." (italics mine, Quine and Ullian 1970, p. 19) Quine elsewhere claims that "We can also *always* get an *absolute* standard by taking in *all* speakers of the language, *or most*." (my italics, Quine 1969, p. 88) (He qualifies this in footnote by disallowing the judgements of "occasional deviants such as the insane or the blind." (Quine 1969, p. 88)) However, this makes what counts as an observation sentence depend on whom we include in the community:

> ...by our definition the observation sentences are the sentences on which all members of the community will agree under uniform stimulation. And what is the criterion of membership in the same community? Simply general fluency of dialogue. This criterion admits of degrees, and indeed we may usefully take the community more narrowly for some studies than for others. What count as observation sentences for a community of specialists would not always so count for a larger community. (Quine 1969, p. 87)

> What counts as an observation sentence will be relative to the community chosen, but this is as it should be; "Deer track" and "Condenser" will qualify as observation sentences for communities of experts and not for wider communities. (Quine and Ullian 1970, p. 19)

Quine claims that we can reduce this community-relativism of observation sentences via a reduction to sentences which serve as observation sentences for the entire linguistic community, although "the practical notion of observation is thus relative to one or another limited community, rather than to the whole speech community." (Quine 1992, p. 6):

> For philosophical purposes we can probe deeper, however, and reach a single standard for the whole speech community. Observable in this sense is whatever would be attested to on the spot by any witness in command of the language and his five senses. If scientists were perversely to persist in demanding further evidence beyond what sufficed for agreement, their observables would reduce for the most part to those of the whole speech community. Just a few, such as the indescribable smell of some uncommon gas, would resist reduction. (Quine 1992, pp. 6-7)

Nonetheless, in the last sentence, Quine hedges his claim of reducibility. It thus seems clear that, for Quine, observation sentences are relative.

For Quine to complete his conceptual programme, he must face the problem of induction. This arises in many guises: determining which stimuli are alike enough to count as identical; deciding how many speakers should agree before a sentence is counted as an observation sentence; and choosing whom to include in the linguistic community. And, of course, the entire problem of determining the words used for agreement and disagreement is a problem of induction (and as noted above, Quine so places it). Also, Quine says of observation sentences that they "can be translated. There is uncertainty, but the situation is the normal inductive one." (Quine 1960, p. 68) Quine concedes that translation requires an inductive method: "As always in radical translation, the starting point is the equating of observation sentences of the two languages by an inductive equating of stimulus meanings." (Quine 1970, p. 179)

Above I argued that Quine attempts to naturalise the conceptual programme by developing an empirical criterion for clarity, where clarity of sentences is measured by the extent to which it commands agreement. This proposal, as I noted, is problematic, since many people agree and have agreed to sentences, for example, about gods, e.g., "God is love" which are surely quite unclear. Further, Quine never specifies how to choose an appropriate linguistic community for the determination of which sentences are observation ones, or a rule for choosing alike enough sets of stimuli. To surmount the last two problems Quine requires a confirmation theory. Quine's theory rests on translation, and translation is, as Quine acknowledges, an inductive process. Thus, for Quine's conceptual programme to proceed, so must the doctrinal programme. Quine does in fact provide an account of how he perceives the doctrinal programme as proceeding, and it is to this I now turn.

G. Quine on the Doctrinal Programme.

Quine regards the process by which a scientist makes an induction as a similar but more complex process of the type by which an infant learns language, which in turn is similar to the way an animal learns. Thus, Quine's account of animal and infant learning are the basis for his account of scientific learning. Just as Quine's account of the conceptual programme is his theory of translation, that is, a scientific approach to meaning, so is his account of the doctrinal programme his theory of learning. I shall examine Quine's learning theory ~~they,~~ and argue that it does not provide the purely empirical foundation for the doctrinal programme. However, I shall also argue that Quine conceives of the doctrinal programme neither as empirical nor apriori, but as conventional. If I am correct, this implies that the standard criticisms of Quine's programme largely miss the point. However, I also argue that his account of the doctrinal programme is incomplete, and that, in particular, it lacks a sufficiently precise theory of confirmation. In the next section I shall attempt to clarify and assess Quine's views on confirmation theory.

1. Canine induction

Quine considers the process by which animals learn to be the simplest possible account of inductive learning, hence I begin with Quine's account of how a dog, to use his example, learns. The learning process can be described in behavioural terms: a dog becomes conditioned to respond to certain stimuli which for it are similar. The conditioned reflex is, according to Quine, inductive learning.

Consider a dog which, on hearing a clatter of pans, runs to the kitchen, apparently expecting food. The dog's expectation, claims Quine, is a conditioned response, a tendency to respond in a stereotyped way to certain stimuli:

> In a dog's experience, a clatter of pans in the kitchen has been followed by something to eat. So now, hearing the clatter again, he goes to the kitchen in expectation of dinner. His going to the kitchen is our evidence of his expectation, if we care to speak of expectation. Or we can skip this intervening variable, as Skinner calls it, and speak merely of reinforced response, conditioned reflex, habit formation. (Quine 1975, p. 69)

A necessary prerequisite for this conditioning, claims Quine, is that the dog group certain kinds of experience together, that is, the noise of the pan on a particular occasion must be grouped with other "similar" noises, and the appearance of food must be likewise grouped with similar experiences:

> When we talk easily of repetition of events, repetition of stimuli, we cover over a certain significant factor. It is the similarity factor. It can be brought into the open by speaking of events rather as unique, dated, unrepeated particulars, and then speaking of similarities between them. Each of the noisy episodes of the pans is a distinct event, however similar, and so is each of the ensuing dinners. What we can say of the dog in those terms is that he hears something similar to the old clatter and proceeds to expect something similar to the old dinner. (Quine 1975, p. 69)

Quine stresses that similarity between events is subjective. (Quine 1975, p. 69) In fact, Quine rejects as meaningless objective notions of similarity, due to the impossibility of defining it in logical or set-theoretic terms. (1987, p. 159-160) But despite its subjective character, Quine finds the notion of similarity important first, because it is an empirical notion in the sense that classes of subjectively similar experiences can be determined empirically, by analysing the dog's behaviour; and second because it helps to explicate induction. As to the first, Quine gives a behaviouristic account of "similarity for the dog":

> We can analyse similarity, for the dog, in terms of his dispositions to behaviour: his patterns of habit formation. His habit of going to the kitchen after a clatter of pans is itself our basis for saying that the clatter events are similar for the dog, and that the dinner events are similar for the dog. It is by experimental reinforcement and extinction along these lines that we can assess similarities for the dog. (Quine 1975, p. 69)

As to the second, according to Quine, induction is the expectation that (subjectively) similar events in the past will be followed by similar events in the future: "...the dog's habit formation, his primitive induction, involved extrapolation along similarity lines: episodes similar to the old clattering episode engendered expectation of episodes similar to the old dinner episode." (Quine 1975, p. 70)

While Quine stresses similarity as the basis of induction, this is only part of the account he need for learning. The other necessary part is expectation. According to Quine, induction is the expectation that subjectively similar events will be followed by some other subjectively similar events. Denote a set of similar of events $\{S_i\}$, and a set of similar responses $\{R_i\}$. In the psychological literature, similarity between events is represented as the probability that the events will elicit a given conditioned response. For example, consider a rat which has been conditioned to press a bar when it hears a 1000Hz tone, so that it always presses the bar when it hears this tone. Suppose that the rat is then exposed to a 999Hz tone. The rat will probably press the bar most of time, but perhaps not all (the same holds for a 1001Hz tone). As the tones presented move further away from 1000Hz, the number of times the response is elicited decreases. The ratio of responses to stimuli can then be represented as a probability distribution. (This procedure is described in Herrnstein 1982.) Quine is presumably proposing that we view induction as a probabilistic function $p(\{S_i\}|\{R_i\})$, successful instances of induction being those matches between stimuli and responses which have probability equal to 1.

This, however, raises the question of how to group certain responses together as relevant for determining what is similar for the subject. We might, speaking behaviouristically, call expectation a conditioned response. We could determine the intensity of this expectation by subjecting animals to various stimuli, and observing their actions. If there are certain stimuli which always yield certain responses, say bells ringing always leads to salivation, we may say that the dog maximally expects to be fed. This, however, might seem to be making the anthropomorphic error of attributing to animals human concerns: we only see stimuli and responses, and it is we who describe certain responses as expectations. However, if we do not anthropomorphise the subject to some degree, the task of uncovering relavant stimulus response pairs would be hopeless. And, the attributions of concerns to the animals seems harmless. Suppose a dog, when subjected to an electric shock (hopefully mild), howls. This does not seem to be an expectation. If, however, a dog drools on hearing a dinner bell, it seems reasonable to describe the dog as expecting food. Presumably the difference is that we have trained the dog so that it drools when it hears the dinner bell, we associate drooling with hunger, and since we know the history of the training, and the nature of the physiological responses, we can just identify certain stimuli/response pairs as expectation. I believe that this account reasonably, if somewhat speculatively, captures Quine's views.

Having settled to his satisfaction the basic idea of induction, Quine goes on to claim that, in the case of a dog, we can determine what a successful induction is, and can then hope to determine why subjective similarity standards should lead to successful inductions: "And now the crux of the problem is the subjectivity of similarity. Why should nature, however lawful, match up at all with the dog's subjective similarity ratings? Here, at its most primitive, is the question 'why is science so successful?'" (Quine 1975, p. 70)

Quine holds that the successful inductions are explained by the natural sciences, in particular, evolutionary biology. According to Quine, evolutionary biology demonstrates that the dog has evolved in such a way that its subjective similarity standards (combined with the appropriate responses, presumably) lead to fulfilled expectations—otherwise its ancestors would not have been successful at reproducing:

> Why should the dog's implicit similarity ratings tend to fit world trends, in such a way as to favour the dog's implicit expectations? An answer is offered by Darwin's theory of natural selection. Individuals whose similarity groupings conduce largely to true expectations have a good chance of finding food and avoiding predators, and so a good chance of living to reproduce their kind. (Quine 1975, p. 70)

Dog learning is, for Quine, the same as human learning: "What I have said of the dog holds equally of us, at least in our pursuit of the rudimentary science of common sense." (ibid) To recap, Quine claims that an animal learns by first grouping certain sets of stimuli together as similar. The animal then uses these as the basis for creating inductive hypotheses, and these hypotheses are frequently confirmed. And the success of these inductions is explained by evolutionary biology.

2. Infant induction

According to Quine, an infant learns pretty much like a dog, by a process of induction over sets of similar stimuli. However, the infant also learns to use language by this process, through making inductive associations of certain stimuli with certain utterances. While the dog reasons inductively about bells and food, the infant makes inductions about words and situations: "...the ostensive learning of words is an implicit case of induction. Implicitly the learner of 'yellow' is working inductively toward a general law of English verbal behavior, though a law he will never state; he is working up to where he can in general judge when an English speaker would assent to 'yellow' or not." (Quine 1969a, p. 12)

Quine claims that the child learns observation sentences first because they are the easiest parts of language to associate with stimuli, and so are the easiest to learn by having queries concerning their use (say, "red?") assented to or dissented from, due to their intersubjective nature:

> At the primitive level, an observation sentence is apt to take the form of a single word, thus 'ball', or 'red'. What makes it easy to learn is the intersubjective observability of the relevant circumstances at the time of utterance. The parent can verify that the child is seeing red at the time, and so can reward the child's assent to the query. Also the child can verify that the parent is seeing red when the parent assents to the query. (Quine 1975, p. 73)

> Primitive inductive learning is evident in the acquisition of various observation sentences. To acquire an observation sentence is to learn when to expect a veteran speaker to approve one's utterance of it, or to assent to it on his own account. This can be learned from sample instances by induction: by extrapolating to further cases along lines of subjective similarity. (Quine 1977 [1979], p. 157)

The account is incomplete, in that we are not told how signs of assent and dissent, the knowledge of which, according to Quine, is fundamental to the acquisition of language, are learned. Also, Quine gives no account of induction: he merely presents the arguments used in an inductive function, leaving that function unspecified.

The importance of similarity standards, for Quine's theory, is clear in the case of infant language learning. Since there are many stimuli associated with any utterance of a word, if Quine's theory is correct, there must be "substantial agreement" as regards similarity between adult and infant:

> In this habit formation the child is in effect determining, by induction, the range of situations in which the adult will assent to the query 'red', or approve the child's utterance of 'red'. He is extrapolating along similarity lines; this red episode is similar to that red episode by his lights. His success depends, therefore, on substantial agreement between his similarity standards and those of the adult... If substantial agreement in similarity standards were not there, this first step in language acquisition would be blocked. (Quine 1975, p. 73)

According to Quine, our similarity standards are, like those of a dog, empirically determinable, although Quine gives no explicit account of how these standards are to be determined. Quine also believes that, our inductions are, in general, successful. From this, he seems to draw the conclusion that our similarity standards do in fact have an objective counterpart:

> In our expectation that subjectively similar events will have similar sequels, we and other animals are often deceived... What is uncanny, however, is how overwhelmingly much more often our expectations are fulfilled than disappointed. We take their fulfillment hour in and hour out as a humdrum matter of course; the occasional unexpectedness is what we notice. Our standard of similarity, for all its subjectivity, is remarkable attuned to the course of nature. For all its subjectivity, in short, it is remarkably objective. (Quine 1987,

p. 160)

The inductive process that operates for our general learning is particularly successful when applied to learning observation sentences, according to Quine: "These linguistic inductions tend to be highly successful—more so than the general run of inductions in our fairly friendly world." (Quine 1977 [1979], p. 157) This, Quine claims, is because our similarity standards concerning words and situations are the product of a common "heredity, environment, and social interaction..." (ibid, p. 158): "Happily the agreement holds; and no wonder, since our similarity standards are a matter partly of natural selection and partly of subsequent experience in a shared environment." (Quine 1975, p. 73) Thus Quine holds that evolution explains why we are so good at making inductions:

> In the light of Darwin's theory of natural selection we can see why this might be. Veridical expectation has survival value in the wild. Innate standards of subjective similarity that promote successful expectation will tend to be handed down through the survival of the fittest. The tendency will have favored us and other species as well. These considerations offer no promise of future success if nature takes what we would regard as a sudden turn, but they do account plausibly for how well we have been doing up to now. (Quine 1987, pp. 160-161)

This concludes my exposition of Quine's theory of infant learning. As is clear, it is basically the same as his account of animal learning, except the infant learns the (stimulus) meanings of words in addition to expecting food after noises emanate from the kitchen. However, we would like to know how many other things are learned, for example, scientific theories. Quine claims to give an empirical account of this, to which I now turn.

3. Scientific learning

Our learning does not end with observation sentences: we must also learn what Quine terms the "superstructure" of language, that is, "past and future tenses,... conditionals and conjecturals and metaphors, and... theoretical and abstract terms." (Quine 1977 [1979], p. 158) It does not seem that the simple theory explicated in the previous two sections can account for the whole of language acquisition. Quine agrees: "Direct conditioning or simple induction does not suffice for the acquisition of language generally." (ibid) Nonetheless, Quine asserts that "It is evident that these further linguistic structures are based, however precariously, on the observational vocabulary that was learned by direct confrontation and simple conditioning..." (ibid) Learning to use these complex grammatical constructions "is beyond the reach of simple induction; it proceeds by imitation and analogy in more complicated ways." (ibid, p. 159)

To understand what Quine means by "simple induction" we must examine his theory of prediction. According to Quine, the most basic predictions are given by what he calls the 'observation categorical,' which is a statement of the form "Whenever this, that", where "this" and "that" are observation sentences. According to Quine, the observation categorical "is the first step in scientific theory": it is scientific because "[i]t is a hypothesis that generates predictions and can be refuted by failure of prediction" (Quine 1987, p. 161); it is a first step because it is the simplest, or most basic, form a hypothesis can take, since it is composed out of the most basic units of language, and it contains no theoretical terms. The use of observation categoricals for the purpose of prediction is what Quine calls "simple induction."

"Sophisticated science far transcends simple induction" (Quine 1987, p. 161) in that it refers to theoretical, as opposed to observation, terms. An example of a theoretical term is one which refers to a body existing through time, such as "apple": "Various of the one-word observation sentences like 'Rabbit' and 'Apple', which were themselves learned in the simple inductive way, will now spawn terms in their likeness—terms denoting bodies. The terms are already theoretical. A body is conceived as retaining its identity over time between appearances." (Quine 1977 [1979], p. 159) Talk of bodies persisting over time, as of other theoretical entities, is carried out in the "superstructure" of language: "Whether we encounter the same body the next time around, the same apple, for instance, or only another one like it, is a question not to be settled by simple induction. It is settled, if at all, by inference from a network of hypotheses that we have internalized little by little in the course of acquiring the nonobservational superstructure of our language." (Quine 1977 [1979], p. 159) Since these hypotheses are framed in a language that is not learned by direct conditioning, their proper use cannot be determined by simple induction. Instead, they must be learned by "the hypothetico-deductive method." According to Quine, the sentences which compose science imply observation categoricals, which are then tested against experience: this defines what Quine means by "the hypothetico-deductive method." (Quine 1987, p. 161):

> These hypotheses are supported only indirectly by past observation: they owe their plausibility to our having inferred other consequences from them that were borne out by observation. Such is the continuing method of science: not simple induction, but the hypothetico-deductive method. (Quine 1977 [1979], p. 159)

> Science departs from simple induction. Science is a ponderous linguistic structure, fabricated of theoretical terms linked by fabricated hypotheses, and keyed to observable events here and there. Indirectly, via this labyrinthine superstructure, the scientist predicts future observations on the basis of past ones; and he may revise the superstructure when the predictions fail. It is no longer simple induction. It is the hypothetico-deductive method. (Quine 1975,

pp. 71-72)

The application of this method, Quine claims, gives rise to "a powerful and virtually conclusive refinement of induction, the method of which John Stuart Mill called concomitant variation" (Quine is referring here to the method of varying the conditions under which a hypothesis is tested.) (Quine 1987, p. 161)

Thus Quine identifies the process of learning a natural language with developing a scientific theory, and claims that an account of language learning will give an account of "scientific evidence":

> There is the beginning, here, of a partnership between the theory of language learning and the theory of scientific evidence.... The channels by which, having learned observation sentences, we acquire theoretical language, are the very channels by which observation lends evidence to scientific theory. It all stands to reason; for language is man-made and the locutions of scientific theory have no meaning but what they acquired by our learning to use them. (Quine 1975, p. 74)

It might seem that Quine is proposing that the conceptual programme be extended to examining not only the meaning of scientific terms, but also the meaning of the word "evidence," as used in science. However, it is also possible to read Quine as claiming that an empirical investigation of what is in fact counted as evidence for a theory will yield some type of normative gain. Consider the following, in which Quine links the study of language learning to the doctrinal programme, and claims that we can give an empirical account of "the relation of evidential support":

> We see, then, a strategy for investigating the relation of evidential support between observation and scientific theory. We can adopt a genetic approach, studying how theoretical language is learned. For the evidential relation is virtually enacted, it would seem, in the learning. This genetic strategy is attractive because the learning of language goes on in the world and is open to scientific study. It is a strategy for the scientific study of scientific method and evidence. We have here a good reason to regard the theory of language as vital to the theory of knowledge. (Quine 1975, pp. 74-75)

While Quine says that we have "good reason" for considering the empirical theory of language learning important to an account of justification, it is not clear if this is because we do happen to in fact find the correct evidential relations when we learn language, 'correct' being defined independently of the theory of language learning; or that there simply is no more to an evidential relation being 'correct' than its being one which we use in learning language.

Elsewhere, Quine says that the observation sentence is linked to the "traditional" doctrinal programme, indicating his aim to produce an empirical doctrinal programme as a counterpart to his empirical conceptual programme:

> [The observation sentence's] relation to doctrine, to our knowledge of what is true, is very much the traditional one: observation sentences are the repository of evidence for scientific hypotheses. (Quine 1969, p. 88)

> To philosophers 'observation sentence' suggests the datum sentences of science. On this score our version is not amiss; for the observation sentences as we have identified them are just the occasion sentences on which there is pretty sure to be firm agreement on the part of well-placed observers. Thus they are just the sentences on which a scientist will tend to fall back when pressed by doubting colleagues. (Quine 1960, p. 44)

These quotes, however, make clear the difficulties of interpreting what Quine says: when he speaks of evidential relations it is not clear what status he assigns them, whether he holds them to be justifiable by empirical means, or whether the practice of language learning happens to coincide with correct inductive practice. I argued in the chapter preceding that a purely empirical account of justification faces insuperable difficulties. However, it is also clear that Quine wishes to avoid any notion of apriori justification. In the next section, I shall argue that in fact Quine occupies a middle position, and a more defensible one, between these two.

4. The doctrinal programme conventionalised

In this section I venture a hypothesis: The foundation of Quine's epistemology is that he takes some form of empiricism is correct, and that this empiricism is accepted on a conventional basis, and not because he believes it to have been proved true apriori or probable aposteriori. However, there is, according to Quine, very little to disagree about once the convention is accepted. I begin by attempting to explain what Quine means by empiricism, and argue that it is much like Berlin's characterisation of empiricism in the motto of this chapter. Quine has stated that he wishes to retain "two cardinal tenets of empiricism": that only sensory evidence counts as evidence for scientific theories, and that the learning of the meanings of words "must rest ultimately on sensory evidence." (Quine 1969, p. 75) I believe that Quine's strategy in implementing his empiricism is to show how what he calls merely "science", which I presume means all "scientific" theories, conforms to certain empirical standards. (Elsewhere, Quine speaks of the empiricist tradition as equivalent to the scientific tradition: "the empiricists' insistence on exclusively sensory evidence is itself scientific." (Quine 1992, p. 8)) Quine accepts an empirical theory of learning, and on the basis of this theory aims to show that this theory supports the rest of science. I believe that this interpretation fits well with the quotation of Quine on circular reasoning at the beginning of this chapter, and that it makes sense of the following quotes:

> I am not appealing to Darwinian biology to justify induction. This would be

> circular, since biological knowledge depends on induction. Rather I am granting the efficacy of induction, and then observing that Darwinian biology, if true, helps explain why induction is as efficacious as it is. (Quine 1975, p. 50)
>
> By sensory evidence I mean stimulation of sensory receptors. I accept our prevailing physical theory and therewith the physiology of my receptors, and then proceed to speculate on how this sensory input supports the very physical theory that I am accepting. I do not claim thereby to be proving the physical theory, so there is no vicious circle. (Quine 1981, p. 24)
>
> ...I shall not be impressed by protests that I am using inductive generalizations, Darwin's and others, to justify induction, and thus reasoning in a circle. The reason I shall not be impressed by this is that my position is a naturalistic one; I see philosophy not as an a priori propaedeutic or groundwork for science, but as continuous with science. I see philosophy and science as in the same boat—a boat which, to revert to Neurath's figure as I so often do, we can rebuild only at sea while staying afloat in it. There is no external vantage point, no first philosophy. (Quine 1969a, p. 14)

I take it that "granting the efficacy of induction", "accept[ing] our prevailing physical theory", and the metaphor of rebuilding the boat in which we are floating all amount to the same thing: assuming the efficacy of the inductive standards of science as it is in fact practised. (Also, Quine's rejection of "first philosophy" in this context perhaps refers to a rejection of metaphysics.) And, as I have shown, Quine's theory of the inductive standards of science lies in his theory of learning, outlined above.

I also find support for my interpretation in Quine's charaterisation of normative reasoning. Quine explicitly claims that it consists solely in ends-means reasoning: that is, normative statements are properly construed as statements of the form "If you want X, then you should do Y." Quine speaks of normative epistemology as a type of engineering in that it employs this means-ends reasoning: "Insofar as theoretical epistemology gets naturalized into a chapter of theoretical science, so normative epistemology gets naturalized into a chapter of engineering: the technology of anticipating sensory stimulation." (Quine 1992, p. 19) Quine has in fact made clear his view that epistemological reasoning can only be of the ends-means kind:

> A word now about the status, for me, of epistemic values. Naturalization of epistemology does not jettison the normative and settle for the indiscriminate description of ongoing procedures. For me normative epistemology is a branch of engineering. It is the technology of truth-seeking, or, in more cautiously epistemic terms, prediction. Like any technology, it makes free use of whatever scientific findings may suit its purpose. It draws upon mathematics in computing standard deviation and probable error and in scouting the gambler's fallacy. It draws upon experimental psychology in exposing perceptual illusions, and upon cognitive psychology in scouting wishful thinking. It draws upon neurology and physics, in a general way, in discounting testimony from occult or

> parapsychological sources. There is no question here of ultimate value, as in morals; it is a matter of efficacy for an ulterior end, truth or prediction. The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter has been expressed. (Quine 1986, pp. 664-665)

Presumably, Quine is attempting to provide hypothetical imperatives of the form "If you want to be an empiricist, then you should do Y." The empiricism is accepted as a convention, not as a well-confirmed or intuitively justified theory.

Talk of conventionally accepting empiricism is rather vague. The next sections will be devoted to sketching how Quine conceives of the empirical convention, and how we accept this convention. Whatever the faults of this programme, and I shall argue that there are several, if I am correct in my interpretation of Quine, the arguments against his programme discussed in the last section miss the point. I will therefore devote the next section to examining these arguments in light of this interpretation.

H. Quine's programmes and the dilemma of naturalism

In this section I will argue that the charges of circularity and self-refutation do not, as I interpret him, carry weight. Consider Quine's views on "simple induction." He assumes that we, for the most part, make successful simple inductions, and that they support certain scientific theories, which are successful. Among these theories is evolutionary biology. This successful theory confirms that in fact we are making successful inductions, because we are well adapted to our world by the processes of natural selection (in that species which do not follow induction become extinct). However, we cannot in any sense prove that simple induction is correct—it is merely accepted by Quine because it is so firmly established as the way we think both in language acquisition and science generally. Given that the epistemic status of "simple induction" is that of a convention, it follows, I shall argue, that Quine's account of the doctrinal programme is neither circularly justified nor self-refuting. I shall, however, argue that Quine assumes too much agreement on what he sees as the conventions underlying science. Accepting the disunity of science leads to a much more pluralistic conception of methodology.

It may appear that assuming scientific theories in order to show that if they were true, then their truth would receive support from their findings is circular. I do not believe that it is. Consider the most famous circular argument in philosophy, the Cartesian circle. The textbook account of this circle is as follows: Descartes argued from "Everything which is clearly and distinctly perceived is true" to "God exists." He then defended this criterion of truth by appeal to God's existence, for if there were a God, he would not deceive us into believing clear and distinct ideas. This is clearly fallacious: taking A to be "Everything which is clear and distinctly perceived is true", and B to be "God exists", the argument is of the form

$$\begin{array}{c} A \vdash B \\ B \vdash A \\ \therefore \vdash A, B \end{array}$$

Both A and B could be false, the premisses could be true and the conclusion false, and hence the argument is invalid. If we were to substitute Behavioural theories of language learning for A, and Evolutionary biology for B, we would likewise get a fallacious argument, and this seems to be the nub of the standard objection to Quine's naturalism.

But Quine need not be interpreted in this manner. Instead, he could, and sometimes seems at times to argue as follows:

$$\begin{array}{c} A \vdash B \\ B \vdash A \\ \therefore \vdash A \leftrightarrow B \end{array}$$

That is, Quine argues that his theory of learning supports evolutionary theory because, if true, it explain how evolutionary theory has gained support from observation, and further that if evolutionary theory is true, that it supports his theory of learning. (I am not here defending Quine's contention that these two theories do support one another: in fact, I doubt it.) Such as argument is clearly valid, but is, however, not very useful or enlightening. Take A to be "The moon is made of green cheese" and B to be "mice would be maximally happy living on the moon." It could be that these two statements support each other: this is of little importance in science, however, since it tells us nothing about the moon/mouse relationship, much less about the composition of the moon.

What is missing in the deductive account of conventionalist reasoning is a mechanism by which theories can be supported by empirical data and in turn support each other on the basis of this data. This can be supplied, however, by a probabilistic account: if A probabilistically supports B, then B probabilistically supports A. I begin with the probabilistic analogue of the deductive circular argument just presented. Obviously:

$$\frac{p(A \wedge B)}{p(B)} > p(A) \leftrightarrow \frac{p(B \wedge A)}{p(A)} > p(B)$$

By the definition of conditional probability, then,

$$p(A|B) > p(A) \leftrightarrow p(B|A) > p(B)$$

That is, A receives support from B if and only if B receives support from A. This does not, of course, establish the truth, or even a high probability for A and B, as the following consistent probability assignment shows: $p(A|B) = .6$, $p(B|A) = .8$, $p(A) = .3$, $p(B) = .4$. Like the preceding deductive account of circular reasoning, the probabilistic account is not at all informative. Again, it may be that the moon's being made of cheese is supported by possible behaviour of mice. In that case, the probability of certain mice-behaviour may be raised by the moon's being made of cheese. Neither is very likely. Likewise, it may be the case that while behavioural learning theory and evolution might support one another, they are both improbable.

The above probabilistic reconstruction of Quine's thesis is lacking in two points. First, it does not capture Quine's claim to be "granting the efficacy of induction", that is, what is conventional; second, it does not take into account any empirical evidence which might be available for his learning theory or evolutionary theory. It might seem that, taking A to be "learning theory" and B to be "evolutionary theory", that Quine is claiming that $p(A) = 1$ (since he is assuming it). But then, B would not be able to provide support for A, since A would already have maximum probability.

Instead, if I am correct in my claim that Quine's theory is best conceived as probabilistic, then we can interpret Quine's granting the efficacy of induction to mean his accepting a probabilistic framework. This probabilistic framework is the same as his theory of learning, and this theory may be checked against empirical evidence using probabilistic induction. One scientific theory which provides this evidence is evolutionary theory. I propose, then, to reconstruct Quine's claim as follows. First, we may reason probabilistically (by assumption). Second, evolutionary theory supports behavioural learning theory, that is,

$$p(A|B) \geq p(A)$$

Finally, there is some evidence which supports B, E:

$$p(B|E) \geq p(B)$$

This, however, is still consistant with both theories having a low probability, and retaining a low probability (if, for instance, E_1 does not support B very much, or B does not support A very much). The appropriate reply is that it is an empirical matter if the theories support each other, and whether or not they can gain enough support for raise their probabilities sufficiently high.

While it much debated as to whether or not evolutionary theory supports a behaviouristic, probabilistic, account of learning, this is an empirical claim. (See, for example, Gigerenzer 1991 and Cosmides and Tooby, 19XX) I have merely aimed to show that Quine is not in fact arguing in a circle. Instead, if I am correct, he assumes some theory of learning as a basis for empiricism (in other words, the learning theory is accepted conventionally), and then argues that we should check whether that theory is supported by the empirical evidence which it tells us how to gather. In a sense, this amounts to an acceptance of a scientific theory on the meta-level, and then testing an analogue of that theory on the object level. The theory is then assessed, by it's own standards, against empirical evidence.

Thus, if Quine's conventionalism means that he is only out to prove arguments in the form above, then I do not believe that he is vulnerable to the charge of circularity. This charge only arises if one makes the fallacious step of inferring either the truth or high probability of A and B. For similar reasons, I do not believe that Quine's epistemology is vulnerable to the charge of self-refutation. As we noted in the chapter previously, the argument that it is self-refuting to argue for the falsity of apriori statements from apriori grounds has considerable force. However, if I am correct, Quine does not argue from an apriori standpoint for the elimination of an apriori standpoint: he instead conventionally chooses an empirical standpoint, and he attempts to argue wholly within this empirical standpoint. Thus, he avoids refuting himself. However, there does seem to me to be something unsatisfactory in the empirical approach outlined above. It seems to me that this approach assumes far too much agreement on the nature of empiricism. I shall address this worry in the next section.

I. Quine on the unity of science

While I do not believe that Quine's approach falls prey to the standard objections, I would agree that it is in some sense unsatisfactory. Quine's views on scepticism make clear that his position is too strong: he apparently argues that it is impossible to consistently adopt a sceptical position, because an empirical standpoint is forced on us. Quine's proposal does not allow for the possibility that the whole of our scientific world view could be wrong, thus ruling out even the possibility of scepticism. But this is surely wrong, as I shall argue in the next section.

As I showed in the previous chapter, Quine attempts to dismiss scepticism on empirical grounds, arguing that "sceptical doubts are scientific doubts" and that the scientific doubts may be answered (Quine 1975, p. 68). Quine claims that "[t]he basis for scepticism is the awareness of illusion, the discovery that we must not always believe our eyes." (Quine 1975, p. 67) According to Quine, sceptical doubts arise because we realise that our predictions concerning the world can, in such cases as "mirages,... bent sticks in water, ...rainbows, after-images, double images, dreams", be incorrect. Thus the sceptic is characterised as someone who does not believe in the external world constructed in the *Aufbau*, due for example to worries, discussed above, that the medium between eyes and objects may not be homogeneous. According to Quine, modern sceptics proceed in a more complicated fashion: they argue that science, if based solely on sensory evidence, is seriously underdetermined by the data:

> The challenge runs as follows. Science itself teaches that there is no clairvoyance; that the only information that can reach our sensory surfaces from external objects must be limited to two-dimensional optical projections and various impacts of air waves on the eardrums and some gaseous reactions in the nasal passages and a few kindred odds and ends. How, the challenge proceeds, could one hope to find out about that external world from such meager traces? In short, if our science were true, how could we know it. (Quine 1973, p. 2)

If it is the case that the sceptic assumes physics and physiology to reach the, rather unexciting, conclusion that scientific theory is fallible, then any response to such a sceptic may likewise assume these theories without fear of arguing in a circle. If someone asserts that not B follows from A, it is not circular reasoning to show them that this is not so and that in fact B does follow from A. Thus Quine says that "[t]he crucial logical point is that the epistemologist is confronting a challenge to natural science that arises from within natural science." (Quine 1973, p. 2)

Quine's characterisation of the sceptic does not seem correct, however, as I argued in the previous chapter. A sceptic need not and generally does not argue on the basis of scientific evidence that physical objects are not what they seem. The sceptic need merely point out that there are infinitely many hypotheses consistent with any sense data we have. Examples of these hypotheses might be "An evil demon is deceiving me into believing that there is an external world," or "I am a disembodied brain in a vat, being fed sensory information by a wicked scientist." Both these hypotheses can account equally well for any data. (This point is made by Stroud 1984) Thus, the sceptic need not assume any particular scientific theory: in fact, the sceptic need not assume anything; the sceptic need only consider the possibility of such demons.

Quine's discussion of the sceptical argument is puzzling, since he seems to miss the point. However, I believe that a reason why Quine refuses to accept the possibility of scepticism is to be found in his conception of "science." Quine often refers to "science" compendiously as the object of the sceptics' attack and the empiricists' faith, in which, apparently, we all share a common belief. For example, Quine claims that "science is self-conscious common sense": "Scientific neologism is itself just linguistic evolution gone self-conscious, as science is self-conscious common sense. And philosophy in turn, as an effort to get clearer on things, is not to be distinguished in essential points of purpose and method from good and bad science." (Quine 1960, pp. 2-3) Quine also seems to identify natural science, naturalised epistemology, and empiricism: "The most notable norm of naturalized epistemology actually coincides with that of traditional epistemology. It is simply the watchword of empiricism: *nihil in mente quod non prius in sensu*. This is a prime specimen of naturalized epistemology, for it is a finding of natural science itself..." (Quine 1992, p.

19)

Thus Quine conceives of "science" as a very broad endeavour, encompassing all ways of collecting empirical data. I believe that this conception results from a combination of Quine's verification theory of meaning and his holism. If Quine is correct, everything we believe, as long as it is stated in words we understand, is related to empirical evidence in some way, and also to every other belief. Since Quine also believes that science is "common sense", he would perhaps argue that everyone accepts some part of some scientific theory. If his holism is correct, consistency requires that everyone who believes part of science must believe all of science. Quine does not produce this argument, but it seems consistent with his other theses. For example, Quine believes that "everything we ever say about the world" is part of a monolithic "science": "this structure of interconnected sentences is a single connected fabric including all sciences, and indeed anything we ever say about the world..." (Quine 1960, p. 12) Of course, someone could accept nothing, that is, be a sceptic, and hence not believe any part of science. However, Quine is of the view that "no inquiry [is] possible without some conceptual scheme." (Quine 1960, p. 4) This conceptual scheme is clearly scientific for Quine. At least, I believe that is what Quine claims in the following quote, where he seems to claim that given an empirical theory of meaning, the scientific scheme must be accepted:

> So the proposition that external things are ultimately to be known only through their action on our bodies should be taken as one among various coordinate truths, in physics and elsewhere, about initially unquestioned physical things. It qualifies the empirical meaning of our talk of physical things, while not questioning the reference... No inquiry being possible without some conceptual scheme, we may as well retain and use the best one we know—right down to the latest detail of quantum mechanics, if we know it and it matters. (Quine 1960, p. 4)

In the previous quote Quine says that "we may as well retain" the scientific scheme. However, elsewhere Quine seems to believe it unthinkable that the scientific scheme could be rejected. Quine sees the positing of externally existing entities, which for Quine, as noted above, is "rudimentary science", as indispensable, and the discussion of the "motive" of such posits as "senseless":

> Men have believed in something very like our common-sense world of external objects as long, surely, as anything properly describable as language has existed; for the teaching of language and the use of it for communication depend on investing linguistic forms with intersubjectively fixed references. It would be senseless to speak of a motive for this archaic and unconscious posit, but we can significantly speak of its function and its survival value; and in these respects the hypothesis of common-sense external objects is quite like that of molecules and electrons. (Quine 1966 [1976], p. 223)

This is quite similar to his discussion in *Word and Object* of the evidence for the reality of external objects. Quine seems to dismiss the possibility of coherently questioning such evidence:

> On the face of it there is a certain verbal perversity in the idea that ordinary talk of familiar physical things is not in large part understood as it stands, or that the familiar physical things are not real, or that evidence for their reality needs to be uncovered. For surely the key words 'understood', 'real', and 'evidence' here are too ill-defined to stand up under such punishment. We should only be depriving them of the very denotations to which they mainly owe such sense as they make to us. It was a lexicographer, Dr. Johnson, who demonstrated the reality of a stone by kicking it; and to begin with, at least, we have little better to go on than Johnsonian usage. The familiar material objects may not be all that is real, but they are admirable examples. (Quine 1960, p. 3)

In fact, Quine does claim that it is fallacious to question the existence of external objects, because our methods of questioning would be dependent on their existence: "We cannot significantly question the reality of the external world, or deny that there is evidence of external objects in the testimony of our senses; for, to do so is simply to dissociate the terms 'reality' and evidence from the very applications which originally did most to invest those terms with whatever intelligibility they may have for us." (Quine 1966 [1976], p. 229) Quine's arguments, as I have construed them, however, do not rule out the possibility that our speech is only sound and fury, signifying nothing. Yet if I am correct in my interpretation of Quine, he has adopted a convention about learning, and hence about induction, and this convention rules out the possibility of the verification theory of meaning being false.

It seems clear that Quine's thesis that we cannot disagree on the scientific conceptual scheme is too strong. First, as I have argued above, Quine's arguments for holism are not conclusive. Secondly, the thesis that we cannot disagree on 'science' would, for example, rule out any methodological disagreement: if we accept the scientific conceptual scheme, according to Quine, we accept, or should accept, all of it. But consider the dispute between Bayesian and classical statisticians. Surely neither is incoherent: they merely disagree on part of science, i.e., that part which deals with statistical inference. In fact, Quine seems to acknowledge that this view is too strong when he recognised the possibility of non-scientific systems which are equally compatible with the data:

> Might another culture, another species, take a radically different line of scientific development, guided by norms that differ sharply from ours but that are justified by their scientific findings as ours are by ours? And might these people predict as successfully and thrive as well as we? Yes, I think that we must admit this as a possibility in principle; that we must admit it even from the point of view of our own science, which is the only point of view I can offer. I

> should be surprised to see this possibility realized, but I cannot picture a disproof. (Quine 1981, p. 181)

This seems a more reasonable response to the sceptic: it may be true that the world will turn out to be radically different from the way most scientists conceive it to be. However, one must start somewhere, and that somewhere forms a scientific framework. This is, however, a conventionalist response. "Our science" must be assumed to be true, and it is within this, rather vague, framework that Quine sees us working. However, due to their fuzziness, it seems to me that Quine's suggestions are not helpful for methodology. Talk of the whole of our scientific theories seems to me to vague to be helpful. Fortunately, however, Quine has specifically endorsed a probabilistic methodology. In the next section I shall substantiate this claim, and in the following, argue that Quine is wrong about the universal acceptance of this theory.

J. Quine's confirmation theory

Quine speaks of "science" as if all that went by that name were on a par. Yet, he clearly does not believe this: for example, he denies telepathy (Quine 1992, p. 19), even though some self-styled scientists believe in it. Quine thus seems to require a theory of confirmation, that is, he must have a means for choosing or preferring some scientific theories over others. Indeed, Quine identifies himself with those "who look upon philosophy primarily as the theory of knowledge," (Quine 1975, p. 67) and, in his *Philosophy of Logic*, Quine seems to identify the philosophy of induction with the theory of knowledge, and hence with the "main stem" of philosophy: "The philosophy of inductive logic... would be in no way distinguishable from philosophy's main stem, the theory of knowledge." (Quine 1970, p. xi) Thus, Quine seems to want a theory of induction, since it follows that he looks on philosophy as the study of inductive logic.

And, as I argued above, Quine's theory of language learning requires a theory a confirmation. This is evident in his exchange with Noam Chomsky. Chomsky argues on probabilistic grounds that Quine's account of language use is inadequate. According to Chomsky, Quine's view that language is a "complex of present dispositions to verbal behavior, in which speakers of the same language have perforce come to resemble one another" (Chomsky quoting Quine, 1975, p. 310) is in need of a probabilistic interpretation: "Presumably, a complex of dispositions is representable as a set of probabilities for utterances (responses) in certain definable circumstances or situations." (Chomsky 1975, p. 310). However, Chomsky argues, this approach is "untenable", since the probability of producing any particular sentence is "zero": "...assuming 'circumstances' and 'situations' to be defined in terms of objective criteria, as Quine insists, it is surely the case that almost all entries in the situation-response matrix are null. That is, in any objectively definable situation, the probability of my producing any given sentence of English is zero, if probabilities are assessed on empirical grounds." (Chomsky 1975, pp. 310-311) Hence, Chomsky also concludes that the probabilities of producing a sentence in Japanese is the same as my probability of producing a sentence in English (i.e., zero): "...in any event it ["the probability of my producing any given sentence of English"] is not detectably different from the probability of my producing some sentence of, say, Japanese." (Chomsky 1975, p. 311)

Quine's response to this argument is instructive. He does not deny that dispositions to particular types of linguistic behavior should be qualified probabilistically, but disagrees with Chomsky's characterisation of the nature of the probability assignment. In particular, he argues that the probability of a disposition to utter a particular sentence is conditional on very specific circumstances, and is hence not zero:

> I am puzzled by how quickly he [Chomsky] turns his back on the crucial phrase "in certain definable 'circumstances.'" Solubility in water would be a pretty idle disposition if defined in terms of the absolute probability of dissolving, without reference to the circumstance of being in water. Weight would be a pretty idle disposition if defined in terms of the absolute probability of falling, without reference to the circumstance of removal of support. Verbal dispositions would be pretty idle if defined in terms of the absolute probability of utterance out of the blue. I, among others, have talked mainly of verbal dispositions in a very specific circumstance: a questionnaire circumstance, the circumstance of being offered a sentence for assent or dissent or indecision or bizarreness reaction. (Quine 1972, pp. 444-445)

(Chomsky's 1975 is his 1969 John Locke Lecture, hence the seeming contradictions in the dates) It is clear that Quine does in fact intend his account of language learning to be not only fallibilistic, but also probabilistic. What then, we may ask, is Quine's theory of probability?

Quine often mentions probability, even though he never provides a systematic exposition. A catalogue of these would be as boring as it would be uninformative, and I therefore restrict myself to only some of Quine's statements in an attempt to 'rationally reconstruct' his views on probability. I believe that Quine adopts, on different occasions, a logical, a Bayesian, and a frequentist approach, although he does not believe any one conception to be applicable to all situations. My account must be, due to Quine's terseness on the subject, speculative. This terseness is somewhat of a puzzle. As is well known, Quine was a close associate of Carnap. Quine reviewed, in what seems to be considerable detail, an early version Carnap's *Logical Foundations of Probability*, and made detailed comments ~~(Quine and Carnap~~ Creath 1990, pp. 399-402). Also, Quine wrote to Carnap that he was planning a course on Carnap's theory of probability: "...I have read "On inductive logic" and "2 concepts of probability" with much pleasure and appreciation. I am much instructed by them, and am glad you are bringing your powers to bear on this problem. I look forward to your book with much interest, and am struck with the idea that I'd like to make a course of it at Harvard." (Quine and Carnap 1990, p. 384) Moreover, Quine claimed that probability has been successfully explicated: "Successful explications have been found for the concepts of deduction, probability, and computability, to name just three." (Quine and Ullian 1970, p. 66) It is therefore to Quine's views on Carnap's inductive logic that I now turn.

Probability of the Carnapian sort is discussed in his famous "Two Dogmas", and rejected along with the analytic/synthetic distinction. Carnap held that probabilities were analytically determined quantities associated with empirical predicates of a language, thus requiring an analytic/synthetic distinction. If this distinction is untenable, so is Carnap's apriori measure over basic empirical predicates. Quine links the problem of making clear the analytic/synthetic distinction to the problem of drawing a distinction between linguistic and factual components of sentences. He seems to see this in turn as somehow analogous to the problem of developing a theory of confirmation:

> ...I hope we are now impressed with how stubbornly the distinction between analytic and synthetic has resisted any straightforward drawing. I am impressed also, apart from prefabricated examples of black and white balls in an urn, with how baffling the problem has always been of arriving at any explicit theory of the empirical confirmation of a synthetic statement. (Quine 1953 [1980], pp. 41-42)

He then seems to claim that this problem arises because it is not possible to draw a distinction between the linguistic and factual components of a sentence, because the quote continues:

> My present suggestion is that it is nonsense, and the root of much nonsense, to speak of a linguistic component and a factual component in the truth of any individual statement. Taken collectively, science has its double dependence upon language and experience; but this duality is not significantly traceable into the statements of science taken one by one. (Quine 1953 [1980], p. 42)

Thus, it seems that Quine holds that there is no such thing as an apriori justified inductive logic.

However, Quine has co-authored a work in which he seems to endorse a limited application of the principle of indifference (Quine and Ullian, 1970: this work contains Quine's lengthiest discussion of probability). After dividing probabilities into epistemic and objective varieties, Quine and Ullian discuss the epistemic variety in connection with gambling, and endorse a version of the principle of indifference:

> Some philosophers of science have tried to apply numerical probabilities to hypotheses. This move is suggested by the reflection that confirmation is a matter or more and less. In games of chance, indeed, the probability of hypotheses makes good sense; in fact, this is where the calculus of probabilities began. There is a clear reason to assign a probability of 1 in not quite 505 to the hypothesis that your next poker hand will be a pat flush. It is clear what information to use in this computation; for we know what cards are in the deck and that you will receive five, while we do not know the order of the cards nor any other influences. This available information consequently reduces the question to a count of combinations. (Quine and Ullian 1970, p. 69)

However, Quine and Ullian raise the well-known problems associated with the logical conception of probability. First, there is a practical problem of framing the hypothesis and evidence: "In the wider world, however, how could we begin to calculate the probability of a hypothesis—say of the hypothesis that the universe began with a bang, or that Babbitt was the murderer? There would be the problem of cataloguing all relevant information." (Quine and Ullian 1970, p. 69) And, as extensively discussed in Chapter I, there are seemingly insurmountable problems with the Principle of Indifference because it does not seem possible to uniquely decompose an event space: "Also there would be the far greater problem, which seems hopeless on the face of it, of compartmenting all alternative possibilities into what could be viewed as equal bits, preparatory to counting combinations." (Quine and Ullian 1970, p. 69)

Quine (with Ullian) is quite sceptical about the possibility of constructing a calculus that would yield exact values for confirmation of hypotheses given evidence: "For the foreseeable future we can do no better on the whole, regarding the degree of confirmation of our hypotheses, than regard some as better confirmed than others and some as not comparable in those terms at all. This we must all see as the practical situation; some philosophers also see it as the necessary situation." (Quine and Ullian 1970, p. 69-70) Hence, while Quine seems to believe that the (logical) probability of an event may be determined exactly in some cases, he doesn't believe it possible to develop an exact confirmation theory, or at least one based on a principle of indifference.

In the above mentioned book, Quine and Ullian also mention the frequentist variety of probability in the context of epidemiology: if 93% of a population inoculated against a particular disease do not catch that disease, then Quine claims that the vaccine is 93% effective, that is, that the probability that it will be effective is .93:

> Cases do sometimes arise, even outside the gambling hells [sic], where we can make reasonable sense of the probability of a hypothesis. Statistics up to now show, let us suppose, that inoculation against some particular disease has been effective in 93 percent of the cases in which it was used. This makes sense of assigning the probability of 93 percent to the hypothesis that Zee, recently

> inoculated and subsequently exposed, will escape the disease. (Quine and Ullian 1970, p. 70)

The tone of the quote seems to indicate that they does not believe that the relative frequency approach covers all uses of probability.

The possibility of a Bayesian interpretation of various aspects of Quine's work has been mentioned by Brian Skyrms (1980, p. 66-70), and it seems to me that such an interpretation of Quine would be quite congenial to the spirit of his enterprise. Quine does in fact discuss the subjectivist approach to probability, and grants that subjective probabilities do measure degrees of belief: "One way of testing belief, powerful where applicable, is by calling upon the professed believer to put his money where his mouth is. Acceptance of a wager evinces sincerity, and the odds accepted conveniently measure the strength of the belief." (1987, pp. 19-20)

However, Quine seems to reject subjective probability in some cases, again on the grounds of limited application. Quine holds to the behaviouristic interpretation of the Dutch Book argument, in that he believes that the bet would actually involve some willingness to part with some good in exchange for another. He does, however, seem to contradict himself on this matter. He first claims that since a bet on a universal hypothesis cannot be settled, the subjective probability cannot be evinced: "But this method is applicable only in cases where the believed proposition is one that can eventually be decided to the satisfaction of both parties, so that the bet can be settled. It is not applicable to the one about beauty, or about one's Redeemer..." (Quine 1987, p. 20) But immediately following, he claims that bets can be measures of degrees of belief, even though they may not be able to be settled:

> Beliefs do sometimes make good behavioral sense without admitting of wagers. This is true of very theoretical beliefs, having to do, say, with the expanding universe or elementary particles or the dawn of language. The turn that one gives to one's research, and the supporting evidence that one marshals or the corollaries that one derives, are substantial indications that one holds the belief, though it be a belief on which a bet could never be settled. (Quine 1987, p. 20)

However, I do not believe that Quine contradicts himself: he merely requires that degrees of belief must in some way be empirically determinable beliefs. As discussed in Chapter 3, there are well known difficulties in such an approach: for example, a person might be disinclined to gamble, in which case they would not have empirically determinable probabilities. Or, if an exchange of money is actually involved, people may hedge their bets, thus not giving their true probabilities.

I conclude that for Quine there are three different types of probability: logical, subjective, and frequentist, although the logical is not clearly distinguished from the other two. More recently, Quine has mentioned only the subjective and frequentist approach. At the end of a recent article praising the virtues of a purely extensional logic, Quine gives his lengthiest discussion of probability in decades. (His consideration of probability seems somehow motivated by the foregoing discussion of extensional and intensional interpretations of logic, but the link is not clear.)

> Finally, how does probability fare under extensionality? Where the probability is statistical, we can settle for a statement of the statistical basis itself as the ratio of sizes of classes of cases. Where the probability is subjective, i.e. the degree of expectation, we are up against issues of behavioral or physiological analysis of mental behavior, irrespective of extensionality. The neat behavioral measure of a subjective probability is the minimum acceptable odds at a wager. (Quine 1994, p. 150)

These brief remarks exhaust Quine's views on confirmation theory, and they are clearly only of a cursory nature. Why then, if Quine does assign such importance to confirmation, is he so vague as to its exact nature? I believe that the answer lies in the difficulty of determining probabilities of the sort which Quine discusses. For example, subjective degrees of belief are not behaviourally determinable in all cases, and it is surely not possible to apply Carnap's inductive method with any generality. Also, relative frequencies give probabilities only in cases in which we can repeat an experiment. If I am correct, then this is what Quine means by his statements in *Pursuit of Truth* that a disposition statement such as "Sodium chloride dissolves in water" is "accepted as a vague statement of strong probability, open to question only where the improbable counter-instance can be plausibly accounted for." Later in the same work, Quine refers to "vague and uncalibrated probabilities.' (1992, p. 95)

Perhaps Quine has paid little attention to probability because he regards the nature of probabilistic confirmation as uncontroversial. For example, in his paper "The scope and language of science" he repeats his view that "science is itself a continuation of common sense." In particular, Quine claims that the scientist's account of evidence is the man-in-the-street's: "The scientist is indistinguishable from the common man in his sense of evidence, except that the scientist is more careful." Any differences between these two notions of confirmation, Quine claims, can be shown to be superficial: the scientist can always convince the common man his notion of evidence is stronger:

> If the scientist sometimes overrules something which a superstitious layman might have called evidence, this may simply be because the scientist has other and contrary evidence which, if patiently presented to the layman bit by bit, would be conceded superior. Or it may be that the layman suffers from some careless chain of reasoning of his own whereby, long since, he came wrongly to reckon certain types of connection as evidential: wrongly in that a careful

> survey of his own ill-observed and long-forgotten steps would suffice to disabuse his. (A likely example is the "gambler's fallacy"...) (Quine 1966 [1976], p. 233)

However, Quine denies that either the scientist or the layman have "an explicit standard of evidence." But, since we have a primitive notion of confirmation, claims Quine, we may not reject the notion of evidence, and, he also seems to claim, that we therefore cannot reject "science":

> If all discourse is mere response to surface irritation, then by what evidence may one man's projection of a world be said to be sounder than another's? If, as suggested earlier, the terms 'reality' and 'evidence' owe their intelligibility to their application in archaic common sense, why may we not then brush aside the presumptions of science? (Quine 1966 [1976], p. 233)

Quine's answer, as we have seen, is "we may not", because we possess a common conception of evidence. In fact, Quine seems to claim that this conception is contained in "probability and mathematical statistics": "It is clearly true, moreover, that one continually reasons not only in refutation of hypotheses but in support of them. this, however, is a matter of arguing logically or probabilistically from other beliefs already held. It is where the technology of probability and mathematical statistics is brought to bear." (Quine 1992, p. 13)

Quine has not only discussed probabilistic theories of confirmation, he has also discussed falsificationism, but such a discussion is unnecessary for my present point, which is that Quine requires a confirmation theory, that he proposes a theory of confirmation, and that this theory of confirmation, while vague, is probabilistic in nature. I have further argued that this vagueness is due to Quine's belief that there is a core of uncontroversial inductive principles. This last point is the most important, for it is how Quine avoids accepting conventionalism. If there is only one convention to be agreed upon, that of some set of inductive principles, then accepting these principles hardly seems to be conventionalism.

I take issue with Quine as to whether or not there is in fact are uncontroversial inductive principles. For example, I argued in Chapter 3 that the Dutch Book argument is in fact invalid. One need only consult a textbook on Bayesian statistics or methodology to see that probabilistic methodologies are not uncontroversial. It seems, then, that we do in fact have a real choice between methodologies, and that this choice cannot be made on the basis of purely apriori or aposteriori considerations.

K. Conclusion

In this chapter I have argued that Quine does not in fact adopt a purely empirical approach to methodology, as he is often interpreted. Instead, I have argued that Quine argues for a naturalisation of what he calls the conceptual programme, and that this is not controversial. After having presented Quine's theory of language learning, which is the same as his account of scientific learning, I argued that Quine does not in fact argue for a purely empirical approach with respect to what he calls the doctrinal programme, that is, the project of justifying our beliefs. Instead, Quine adopts a more sophisticated conventional cum empirical approach. This approach avoids the dilemma I raised for the empirical approach in chapter 4: that if it argued for from apriori grounds then it is self-refuting, and if it is argued for from aposteriori grounds is circularly justified.

There is much to disagree with in Quine's analysis of epistemology, for example, his account of the failure of the *Aufbau*, his claim that learning is behavioural, and his views on confirmation theory. However, I am, in general, sympathetic to Quine's conventionalism. My main criticism of Quine's conventionally based methodology is that it is too vague to be of much guidance in methodology. Quine assumes too much unity in scientific thought generally, and in inductive methodology in specific. Thus, while Quine espouses empiricism-cum-conventionlism with respect to methodology, he never explores is consequences. In the next chapter I consider the work of Karl Popper, who explicitly adopted conventionalism.

PART III: METHODOLOGY AS CONVENTIONAL

In this final Part I examine conventionalism with respect to methodology. The next chapter is an exposition of Karl Popper's conventionalism. I argue that while Popper has at times accepted a conventionalist account of methodology, he also has endorsed a naturalistic approach. The final chapter is a summary of the results leading to my conclusion that ~~methodology~~ conventionalism is the only viable alternative for methodology.

CHAPTER 7: Popper's Conventionalism

A. Introduction

Karl Popper proposed an alternative to both empirical and apriori approaches to methodology. He argued that methodological statements should be regarded as conventions. In this section I will set out Popper's claim to have established a non-empirical base for methodology, and to have shown that naturalistic approaches to methodology are untenable. I will then turn to criticisms made by Kuhn, Lakatos and Sarkar that Popper's own methodology nevertheless relies in an essential way on appeals to the practice of science (and is thus empirical), and I shall argue that while their conclusions concerning Popper's methodology are correct, their arguments for this conclusion are wide of the mark. I shall then examine Elie Zahar's defense of Popper's methodology against the criticism that it is implicitly naturalistic. I conclude that while Zahar's interpretation is the best that can be given, and while some (or even most) of Popper's methodology can be interpreted as non-empirical, some cannot be so interpreted. Nonetheless, I shall, in the next chapter, conclude in favour of conventionalism.

B. Popper's methodology - Falsificationism

Before I can discuss Popper's arguments against naturalism, I must present a brief summary of those aspects of his methodology which are relevant to the discussion. The cornerstone of Popper's methodology is falsificationism. Popper argues that scientific theories can never be verified, or even confirmed, by observation, but stresses that theories may however be refuted by an application of *modus tollens*: Consider a theory T which predicts the occurrence of event O. Further suppose that O does not happen. By *modus tollens* (if T then O, not O, therefore not T) the theory must be false, that is, it is falsified. According to Popper (and this is the core of falsificationism), a theory is scientific only if it is in principle falsifiable, thus, falsifiability serves as what he calls a demarcation criterion between science and non-science.

C. Popper's argument against naturalism

Popper is highly critical of and has explicitly repudiated naturalism with respect to methodology. In his book *The Logic of Scientific Discovery*, Popper presents his most explicit argument against a naturalistic methodology of science. Due to the centrality of this argument for the position Popper adopts, I shall quote it in full:

> ...*what I call 'methodology' should not be taken for an empirical science.* I do not believe that it is possible to decide, by using the methods of an empirical science, such controversial questions as whether science actually uses a principle of induction or not. And my doubts increase when I remember that what it is to be called a 'science' and who is to be called a 'scientist' must always remain a matter of convention or decision.
>
> I believe that questions of this kind should be treated in a different way. For example, we may consider and compare two different systems of methodological rules; one with, and one without, a principle of induction. And we may then examine whether such a principle, once introduced, can be applied without giving rise to inconsistencies; whether it helps us; and whether we really need it. It is this type of inquiry which leads me to dispense with the principle of induction: not because such a principle is as a matter of fact never used in science, but because I think that it is not needed; that it does not help us; and that it even gives rise to inconsistencies.
>
> *Thus I reject the naturalistic view.* It is uncritical. Its upholders fail to notice that whenever they believe themselves to have discovered a fact, they have only proposed a convention. Hence the convention is liable to turn into a dogma. This criticism of the naturalistic view applies not only to its criterion of meaning, but also to its idea of science, and consequently to its idea of empirical method. (1968, pp. 52-3)

Popper's argument seems to me rather hard to follow, but its aim is clear, which is to justify his rejection of the naturalistic view: "...'methodology' should not be taken for an empirical science". Since this is Popper's most extensively developed argument against the empirical approach to methodology, it is necessary to give it close attention. This will occupy the next four subsections.

1. The Conventionality of Methodology

The first step of Popper's argument is to claim that it is not possible to decide (by examining what scientists do) whether, to take Popper's example, "science actually uses a principle of induction." In Popper's terminology, "A principle of induction would be a statement with the help of which we could put inductive inferences into a logically acceptable form" (1968, p. 28). Presumably Popper means by this that a principle of induction is a formal explication of the notion of inductive inference. According to Popper, "It is usual to call an inference 'inductive' if it passes from *singular statements* (sometimes also called 'particular statements'), such as accounts of the results of observations or experiments, to *universal statements*, such as hypotheses or theories" (1968, p. 27).

Popper's belief that a naturalist approach cannot settle the question of whether or not science uses a principle of induction is, he claims, supported by the fact that the definition of 'science' and 'scientist' is a convention:

> I do not believe that it is possible to decide, by using the methods of an empirical science, such controversial questions as whether science actually uses a principle of induction or not. And my doubts increase when I remember that what it is to be called a 'science' and who is to be called a 'scientist' must always remain a matter of convention or decision. (1968, p. 52)

Popper's observation that definitions of 'scientist' and 'science' are conventional seems crucial to his argument, and one which seems to be linked to his view that methodologies are conventions. He seems to argue that since methodological statements about 'scientists' and any 'principle of induction' are conventions, so too are methodologies. But it is clear that Popper regards methodology as a convention with no apriori basis:

> My criterion of demarcation will accordingly have to be regarded as a *proposal for an agreement or convention*. As to the suitability of any such convention opinions may differ; and a reasonable discussion of these questions is only possible between parties having some purpose in common. *The choice of that purpose must, of course, be ultimately a matter of decision, going beyond rational argument*. (Popper 1968, p. 37, second set of italics mine)

Conventions are, of course, familiar features of the social world. Traffic lights, for example, are usually arranged vertically, with red on top, green on the bottom; but they are sometimes in the United States arranged horizontally. There is no wrong or right way to arrange them, although one of these may recommend itself on grounds of convenience. Since conventions are neither true nor false, it would seem that Popper is claiming the same for methodology.

This might seem a rather disappointing conclusion to reach, and indeed, Popper seems to draw back from it: in a footnote added in a later edition to the above statement concerning the conventionality of methodology, Popper states that "a reasonable discussion is always possible between parties interested in truth, and ready to pay attention to each other." (1968, p. 37) It is not clear what Popper means by "reasonable discussion". Since he holds that methodologies are conventions, and since conventions have no truth value, it is hard to see how "an interest in truth" can adjudicate between rival conventions.

However, Popper holds that methodologies that incorporate a principle of induction are in error:

> Some who believe in inductive logic are anxious to point out, with Reichenbach, that 'the principle of induction is unreservedly accepted by the whole of science and that no man can seriously doubt this principle in everyday life either'. Yet even supposing this were the case—for after all, 'the whole of science' might err—I should still contend that a principle of induction is superfluous, and that it must lead to logical inconsistencies. (1968, p. 29)

This illustrates a tension in Popper's methodology, which I will discuss more fully below. On the one hand, Popper claims that methodologies, and his methodology in particular, are conventions, with the happy consequence that it avoids the difficulties of justifying a methodology *apriori*. On the other hand, it might seem to imply that a choice of a methodology is, from an epistemic point of view, quite arbitrary. This is a position Popper (understandably) also wishes to avoid, which, as we shall see, leads him to reintroduce a non-conventional approach. However, for the purpose of this argument, I shall proceed on the basis of Popper's claim that methodologies are conventions.

2. Non-naturalistic criteria for the assessment of methodologies

The second step in Popper's argument is to show, without appealing to scientific practice, how to determine whether or not "science actually uses a principle of induction." This is to "compare two different systems of methodological rules; one with, and one without, a principle of induction", and then check to see whether, first, the principle makes the system inconsistent, second "whether it helps us", and third "whether we really need it". (Popper holds that a methodology which includes a principle of induction fails on all three counts):

> I believe that questions of this kind ["whether science actually uses a principle of induction or not"] should be treated in a different way [than "using the methods of an empirical science"]. For example, we may consider and compare two different systems of methodological rules; one with, and one without, a principle of induction. And we may then examine whether such a principle, once introduced, can be applied without giving rise to inconsistencies; whether it helps us; and whether we really need it. It is this type of inquiry which leads me to dispense with the principle of induction: not because such a principle is as a matter of fact never used in science, but because I think that it is not needed; that it does not help us; and that it even gives rise to inconsistencies. (1968, p. 52-53)

Popper thus proposes three methodological criteria: consistency, fruitfulness, and indispensability. The criterion of consistency appears to be apriori, and so it might seem that Popper does not view methodology as completely conventional. On the other hand, Popper seems, surprisingly, to defend consistency in a methodology on pragmatic grounds. He argues that since any proposition can be deduced from a contradiction, a theory containing a contradiction "can give us no information at all. A theory which involves a contradiction is therefore entirely useless *as a theory.*" (Popper 1965, p. 319). Popper considers three possible defences of an inconsistent theory: First, it may be "interesting in itself", second, "it may give rise to corrections which make it consistent," and third, we may find a method which allows us to avoid the "false conclusions which are admittedly entailed by the theory" (1965, p. 321).

However, Popper rejects these defences, pointing out that a methodology which allows contradictory theories will not lead to a search for better theories: "...if we seriously intend to put up with it [a contradiction] then nothing will make us search for a better theory." (Popper 1965, p. 321) Thus Popper holds that such a methodology leads to an uncritical attitude:

> The acceptance of contradictions must lead here as everywhere to the end of criticism, and thus to the collapse of science....to say that the contradictions need not be avoided, or perhaps even that they cannot be avoided, must lead to the breakdown of science, and of criticism, i.e. of rationality. (1965, p. 322)

Popper investigated the possibility of admitting contradictions into science by adopting a logic in which not every sentence in a language follows from a contradiction. However, he rejects such logics on the grounds that they are too weak:

> Very few of the ordinary rules of inference are left [in "a system of logic in which contradictory statements do not entail every statement"], not even the *modus ponens*... In my opinion, such a system is of no use for drawing inferences although it may perhaps have some interest for those who are specially interested in the construction of formal systems as such. (1965, p. 321)

So, it seems that Popper adopts the methodological rule to exclude contradictions from science not for apriori reasons connected with questions of truth, but for pragmatic reasons: a contradictory theory is uninformative. Therefore, it seems that Popper's three criteria arise from his conventional conception of methodology, with the conventions chosen for pragmatic purposes.

3. The naturalistic method leads to dogmatism

So Popper first argues that empirical methods cannot settle methodological questions, due to the conventional character of such questions, and second, that we can use conventional criteria to settle them. Finally, Popper argues that naturalists, who model their methodology on empirical science will, as a result, confuse facts with conventions which "is liable to" lead to dogmatism. Popper gives no example of such a confusion nor any illustration of how confusion could arise; perhaps the following illustrates what he had in mind: Suppose someone who adopted the naturalistic approach, after close study of their practices, concluded that scientists all wore white lab coats, and operated expensive and complex machinery. Popper might deny that this is a factual discovery, since it was a conventional decision to group such people together and call them 'scientists.' He seems to claim further that if one treats the outcomes of methodological inquiries as purely factual, ignoring their conventional aspect, then the criteria for evaluating the conventions outlined in his account of the non-naturalistic approach (consistency, fruitfulness and indispensability) will seem irrelevant, and the conventions will then be accepted without examination, i.e., dogmatically:

> Thus I reject the naturalistic view. It is uncritical. Its upholders fail to notice that whenever they believe themselves to have discovered a fact, they have only proposed a convention. Hence the convention is liable to turn into a dogma. (1968, p. 52-3)

To make clear the force of Popper's charge, it is necessary to examine what he means by "dogmatic." Popper holds that falsifiability serves as a criterion to demarcate science from non-science. He holds that if a proposition is unfalsifiable, it can be held on to whatever the empirical evidence, an attitude Popper calls dogmatic:

> ...the distinction between dogmatic and critical thinking, or the dogmatic and the critical attitude, brings us right back to our central problem. For the dogmatic attitude is clearly related to the tendency to *verify* our laws and schemata by seeking to apply them and to confirm them, even to the point of neglecting refutations, whereas the critical attitude is one of readiness to change them—to test them; to refute them; to *falsify* them, if possible. This suggests that we may identify the critical attitude with the scientific attitude, and the dogmatic attitude with the one which we have described as pseudo-scientific. (1965, p. 50)

4. Summary of the argument

To summarise, Popper argues against a naturalistic methodology in three steps: First, any empirical investigation of science requires a definition of 'science' and 'scientist', and this is a convention. Second, such conventions can be assessed according to their consistency, fruitfulness, and whether or not they are indispensable. Third, since they have no pretensions to describing the world, conventions cannot be assessed using the methods of empirical science. Thus, if we attempt to use the methods of empirical science to study science we cannot assess the conventional decisions that form the foundation of our enterprise. From this Popper concludes that an empirical approach must lead to dogmatism. Since dogmatism is evidently a bad thing, we should not adopt the empirical approach to the study of science, but instead adopt a non-naturalistic methodology. Although Popper does not explicitly say so, it seems that his argument, if correct, shows that naturalistic methodologies are dogmatic, and hence non-scientific, and thus self-defeating.

Popper's discussion of the use of probability statements in physics illustrates his conventionalist approach. Popper holds that these probability statements purport to describe objective probabilities, and that the probability of a particular physical event, say, the decay of a radioactive particle, is a physical property of that particle. Thus, according to Popper, probabilities, as used in physics, are not epistemic, and do not represent degrees of belief about a hypothesis: they are physical properties. Popper claims that his views on the use of probability statements accords with their actual use in physics: "I believe that physics uses probability statements only in the way I have discussed..." However, he immediately states that he is not willing to defend this view ("But I should like to decline to join in any dispute about how physicists 'in fact' proceed..."), since it is, according to Popper "largely a matter of interpretation". Popper explicitly rejects an apriorist interpretation of the dispute over the use of probability statements:

> We have here quite a nice illustration of the contrast between my view and what I called, in section 10 [quoted above, section C], the 'naturalistic' view. What can be shown is, first, the internal logical consistency of my view, and secondly, that it is free from those difficulties which beset other views. *Admittedly it is impossible to prove that my view is correct, and a controversy with upholders of another logic of science may well be futile. All that can be shown* is that my approach to this particular problem is a consequence of the conception of science for which I have been arguing. (1968, p. 262, italics mine)

This reinforces the position that Popper views his methodology as neither true nor false, but conventional. Note however that he says that "a controversy...*may* well be futile", not that it *would* necessarily be futile. Popper seems undecided as to the conventionality of methodology. If methodology is conventional, then it would seem that a dispute would be futile, because there can be no epistemic reason for choosing one methodology over another. If a dispute over methodology would not be futile, then it would seem that there are reasons for choosing one method over another. I shall discuss in the next chapter what these reasons might be.

Popper criticises the empirical approach to methodology on the grounds that it leads to dogmatism. I will assess Popper's argument below. But first, I will examine the claims of Kuhn, Sarkar and Lakatos that notwithstanding Popper's rejection of the empirical approach, his methodology does in fact rely on the practice of science for support, and will argue that their arguments are not successful.

D. Unsuccessful arguments

1. Kuhn's interpretation of Popper

The dispute between Popper and Kuhn as to the empirical nature of methodology is found in Lakatos and Musgrave 1970. Kuhn's methodology is avowedly naturalistic: He develops his methodology out of a close study of scientific practice, and defends it by appeal to what he regards as the central features of that practice. However, Kuhn does not regard methodology as either purely descriptive or purely normative, but as a mixture of the two. Thus he says that his methodology

> ...should be read in both ways at once. If I have a theory of how and why science works, it must necessarily have implications for the way in which scientists should behave if their enterprise is to flourish. The structure of my argument is simple and, I think, unexceptionable: scientists behave in the following ways; those modes of behaviour have (here theory enters) the following essential functions; in the absence of an alternate mode *that would serve similar functions*, scientists should behave essentially as they do if their concern is to improve scientific knowledge. (Kuhn 1970, p. 237)

Kuhn argues that despite the partly naturalistic character of his methodology and the claimed non-empirical character of Popper's, in fact the two are similar. Kuhn claims that both he and Popper agree that the development of science constrains the philosophy of science:

> We are both concerned with the dynamic process by which scientific knowledge is acquired rather than with the logical structure of the products of scientific research. Given that concern, both of us emphasize, as legitimate data, the facts and also the spirit of actual scientific life, and both of us turn often to history to find them. From this pool of shared data, we draw many of the same conclusions. (1970, pp. 1-2)
>
> Among the most fundamental issues on which Sir Karl and I agree is our insistence that an analysis of the development of scientific knowledge must take account of the way science has actually been practised. (1970, p. 4)

Where Kuhn sees their positions diverging is in Popper's taking too narrow a view of the history of science and modelling his methodology on the rather exceptional events that Kuhn calls "extraordinary research," rather than on the more common "normal science." Thus Kuhn takes note of Popper's insistence that methodology must account for so-called crucial experiments in which an established theory is entirely overthrown, such as "Lavoisier's experiments on calcination, the eclipse expedition of 1919, and the recent experiments on parity conservation." Kuhn claims that any methodology built on such dramatic episodes alone would be inadequately grounded, as such episodes are "very rare in the development of science" (1970, p. 5). Such an emphasis on the revolutionary episodes in the history of science, Kuhn claims, prevents Popper from producing a correct account of the growth of knowledge: "...neither science nor the development of knowledge is likely to be understood if research is viewed exclusively through the revolutions it occasionally produces" (1970, p. 6). Kuhn suggests that a better insight into the nature of science as it proceeds normally would be afforded by empirical research into the psychology of groups of scientists.

In his response to Kuhn, "Normal Science and its Dangers", Popper disagrees that it is appropriate to appeal to scientific practice in support of a methodology, confirming his anti-naturalistic stance: "...to me the idea of turning for enlightenment concerning the aims of science, and its possible progress, to sociology or to psychology (or...to the history of science) is surprising and disappointing." (1970, p. 57) Popper here offers two arguments for finding naturalistic methodology untenable: first because sociology and psychology are poorly developed, and secondly because he accepts a version of what I have called the 'circle argument' (discussed in Chapter 4).

In his response ("Reflections on My Critics", 1970) Kuhn argues that despite his protestations to the contrary, Popper's own approach belies his anti-naturalist claims, because he often cites the practice of science in his writings:

> If... [Popper] is challenging the relevance to philosophy of science of the sorts of observations collected by historians and sociologists, I wonder how his own work is to be understood. His writings are crowded with historical examples and with generalizations about scientific behaviour, some of them discussed in my earlier essay [recall that Kuhn cited "Lavoisier's experiments on calcination, the eclipse expedition of 1919, and the recent experiments on parity conservation" as examples]. He does write on historical themes, and he cites those papers in his central philosophical works. A consistent interest in historical problems and a willingness to engage in historical problems distinguishes the men he has trained from the members of any other current school in philosophy of science. On these points I am an unrepentant Popperian. (1970, p. 235-6)

Kuhn is certainly correct that Popper does frequently draw on the history of science. But this does not imply that Popper's methodology is an empirical theory. Popper may only be presenting these episodes as interesting examples to illustrate his philosophical theses. Further, Popper could argue that scientists operate according to the procedures he indicates because they agree with him. Popper is not necessarily relying on historic examples for support of his thesis, and thus his use of these examples does not imply that his approach is empirical. I thus conclude that Kuhn has not established that Popper's method is empirical. I shall, however, argue in section G, that it can be established that Popper does at times appeal to scientific practice in support of his method.

2. Sarkar's interpretation of Popper

Husain Sarkar devotes a chapter of his 1983 *A Theory of Method* to "Popper's Theory of Method", where he denies Popper's claim that methodologies are conventional, and argues that methodologies should be regarded as having a normative status:

> ...I shall argue against the cardinal feature of [Popper's] theory [of method]—that statements of a method are statements of conventions—and show that it is not even necessary for Popper to stake such a claim.... I shall argue that when a method is regarded as empirical or conventional we are led to intractable problems, but that when a method is regarded as normative, with an attendant truth-value, various pieces of the puzzle seem to fall into place. (Sarkar 1983, p. 29)

Sarkar, like Kuhn, tries to show that Popper's methodology is an empirical theory. He bases this charge on Popper's discussion of dialectic (in the essay "What is Dialectic?" (1965)), where Popper claims that dialectic is a type of trial and error, which for Popper is the proper scientific method. Popper holds some episodes in "the history of thought" to be dialectical in this sense, and that dialectic is an "empirical descriptive theory." Sarkar claims that since for Popper dialectic is the proper scientific method, and describes "fairly well" episodes in the history of science, Popper must hold that scientific method is descriptive, and thus not conventional:

> According to Popper, scientific method is a trial-and-error process carried out consciously and systematically; but, he says, this trial-and-error method is nothing but a good description of human activity in general and scientific activity in particular... It is of interest to note that Popper concludes that "Dialectic...is, therefore, an empirical descriptive theory," *because* the method of trial-and-error is. Conclusion: Isn't Popper's method naturalistic (at least in part)? If not, how could conventions *describe*? (Sarkar 1983, pp. 48-49)

Indeed, in his essay, Popper is clearly treating dialectic as an empirical thesis, and so it seems reasonable to assume that Popper is discussing the theory of dialectic as an empirical, and not a normative, theory:

> Dialectic, or more precisely, the theory of the dialectic triad [thesis, antithesis, synthesis], maintains that certain developments, or certain historical processes, occur in a certain typical way. It is, therefore, an empirical descriptive theory, comparable, for instance, with the theory which maintains that most living organisms increase their size during some stage of their development, then remain constant, and finally decrease until they die; or with the theory which maintains that opinions are held first dogmatically, then sceptically, and only afterwards, in a third stage, in a scientific, i.e. critical, spirit. (1965, p. 322)

Sarkar clearly means to criticise Popper on the grounds that his methodology must be at least partly empirical, since it, according to Popper, can serve as an empirical theory. Sarkar's criticism is similar to Kuhn's, and shares its weakness. Popper could respond to Sarkar as he might have done to Kuhn: trial-and-error may provide an adequate description of how science actually proceeds. But this does not imply that this method is nothing but an empirical theory. Popper could argue in two ways. The first would be to abandon conventionalism, and respond that his view of methodology is correct, drawing its support from apriori grounds, and that by reflecting scientific practice merely attests to the rationality of many leading scientists. Popper's theory can be apriori, and yet correspond closely to the facts of scientific practice, much as an ethical theory might correspond closely to the practice of nuns in a nunnery. This, however, would not be in keeping with Popper's conventionalism. Instead, it could be argued that it may just turn out that the conventions Popper proposes happen to be very much like those used by working scientists. In fact, this may be why those conventions were chosen. This does not, however, imply that the conventions governing scientific method are the same as the empirical thesis describing that method. I conclude that Sarkar's, like Kuhn's, arguments that Popper's methodology is naturalistic are inconclusive.

3. Lakatos's interpretation of Popper

Lakatos, like Kuhn and Sarkar, claims that Popper adopts the empirical approach. Lakatos first interprets Popper as arguing for a 'Euclidean' method, that is, one in which methodological rules are *apriori*:

> 'Euclidean' methodologies lay down *a priori general rules* for scientific appraisal. This approach is most powerfully represented today by Popper. In Popper's view there must be the constitutional authority of an *immutable statute law* (laid down in his demarcation criterion) to distinguish between good and bad science. (Lakatos 1968, p. 121)

But Lakatos notes that because Popper arrived at his methodology through consideration of a variety of theories from physics, psychology and economics, he can be seen not as an apriorist but as a naturalist: "After all, Popper defined 'science' in such a way that it should include the refuted Newtonian theory and exclude unrefuted astrology, Marxism and Freudianism." Indeed, Popper's autobiographical account of the development of his methodology seems to bear out Lakatos's claim. Popper claims to have come to his methodology by considering what he considered to be inadequate theories such as "the Marxist theory of history, psycho-analysis, and individual psychology" as opposed to "physical theories,... Newton's theory, and especially... the theory of relativity..." (Popper 1965, p. 34)

But again, although this is suggestive, it is not decisive in establishing that Popper's approach to methodology is an empirical theory. Popper's motives in formulating his methodology are independent of any justification of that method. Just because these episodes in the history of science inspired Popper to formulate his methodology does not mean that he looks to, or needs to look to, these episodes for justification, or that the methodology he arrived at is necessarily an empirical theory.

E. Zahar's interpretation of Popper

Popper has, however, made statements which if taken at face value flatly contradict his account of methodology as non-empirical. For example, as Lakatos points out, later in Popper's essay discussed above, Popper tells us that he rejects the verificationist theory of meaning as a methodological principle because it does not capture actual scientific practice. The verifiability criterion, Popper says "is too narrow (*and* too wide): it excludes from science practically everything that is, in fact, characteristic of it (while failing to exclude astrology)." (Popper 1965, p. 40) And at the end of the essay "What is Dialectic" discussed by Sarkar, Popper seems to claim that philosophers should study methodology in an empirical manner:

> The whole development of dialectic should be a warning against the dangers inherent in philosophical system-building. It should remind us that philosophy must not be made a basis for any sort of scientific system and that philosophers should be much more modest in their claims. One task which they can fulfil quite usefully is the study of the critical methods of science. (Popper 1965, p. 335)

In *The Logic of Scientific Discovery*, Popper concedes that a naturalistic approach may be valuable for philosophy of science (though he does not say wherein its value lies): "A naturalistic methodology (sometimes called an 'inductive theory of science') has its value, no doubt. A student of the logic of science may well take an interest in it, and learn from it." (Popper 1968, p. 52) In *Realism and the Aim of Science* (1983), Popper says that "the subjectivist theory of knowledge, and with it the instrumentalist interpretation of scientific theories, clashes not only with common sense but also with science, and with the rationalist tradition". Popper concludes that therefore it "may be rejected." (Popper 1983, p. 131)

Elie Zahar points to a similar passage in *Conjectures and Refutations*, where Popper rejected instrumentalism because it misrepresents the toals of scientists:

> ...instrumentalism is unable to account for the importance to pure science of testing severely even the most remote implications of its theories, since it is unable to account for the pure scientist's interest in truth and falsity. In contrast to the highly critical attitude requisite in the pure scientist, the attitude of instrumentalism (like that of applied science) is one of complacency at the success of applications. (Popper 1965, p. 114)

Popper concludes that instrumentalism cannot fully capture scientific progress in the way that his methodology can. For Popper, theory precedes observation, that is, theories lead to new observations, but new observations do not lead to new theories: "It is my belief that our discoveries are guided by theory in these as in most other cases, rather than that theories are the result of discoveries 'due to observation'; for observation itself tends to be guided by theory." Popper takes as an example geography: "Even geographical discoveries

(Columbus, Franklin, the two Nordenskjölds, Nansen, Wegener, and Heyerdahl's Kon-Tiki expedition) are often undertaken with the aim of testing a theory." Popper holds that instrumentalism cannot account for these episodes, and clearly takes this inability as a mark against instrumentalism: "Not to be content with offering predictions, but to create new situations for new kinds of tests: this is a function of theories which instrumentalism can hardly explain without surrendering its main tenets." (1965, p. 118) It is difficult to interpret Popper's discussion in any other way except as saying that instrumentalism cannot capture important parts of scientific practice, and for that reason should be rejected, and that his methodology, by contrast, can capture these parts, and should therefore be accepted.

So, on the one hand, Popper denies that his methodology is a descriptive theory, and on the other hand, he rejects alternative methodologies because they fail to explain scientific behaviour and commends his own because it does. Popper cannot respond that his critics have confused the psychological origins of his theory with its justification, or that there is a happy overlap between his non-empirical theory and scientific practice, because he is explicitly appealing to scientific practice to reject the verifiability criterion and instrumentalism.

1. Popper's "quasi-empirical method"

Zahar, in his review of Popper's *Die Beiden Grundprobleme der Erkenntnistheorie* (written in 1930/31 but published only in 1979), attempts to reconcile these two strains in Popper's methodological writings. Zahar defends Popper's claim that his methodology is not an empirical theory, and argues that it only appears to be so because Popper uses the term "science" not in a straightforwardly descriptive manner, but as a commendatory term; according to Zahar, in the contexts where science is appealed to by Popper, "science" should be read as "good science." When Popper says that instrumentalism or the verifiability criterion of meaning should be rejected because it does not square with scientific practice, he means that they do not square with *good* scientific practice. Scientific practice is not labelled "good" on empirical grounds, and so Popper's methodology makes no essential use of actual practice. Zahar refers to this approach as "quasi-empirical."

According to Zahar, Popper, in *Die Beiden Grundprobleme*, proposed a "method of epistemological criticism...which he terms transcendental." Zahar tells us that this method is "immanent", and has "two components: a logical component using the principle of non-contradiction and a quasi-empirical component which consists in confronting a given methodology with actual scientific practice." (Zahar 1983, p. 150-151) Hence Popper's theory of knowledge seems to be a science, thereby making epistemology naturalistic:

> The theory of knowledge is a science of science (*Wissenschaftswissenschaft*). It behaves with respect to the sciences in the same way as the latter behave with regard to experience. The transcendental method is an analogue of the empirical method. The theory of knowledge is thus a theoretical science, which of course contains certain stipulations (definitions for example). It does not however consist only of arbitrary conventions but also of propositions which can be refuted through a comparison with the methods which are effectively and successfully used with the sciences. (Popper, Zahar's translation, 1983 p. 151)

This seems an unequivocal statement of the naturalist position, in sharp contrast with the diametrically opposite view Popper takes in *The Logic of Scientific Discovery*. Did Popper fail to notice the contradiction in his position, or did he simply change his mind over time without troubling to refer back to his earlier views? It would seem that in *Die Beiden Grundprobleme* (1930/31) he espoused a naturalistic methodology, in *The Logic of Scientific Discovery* (1934) he rejected such a view, and in *Conjectures and Refutations* (the essays quoted are from 1953, 1956, and 1940) he once again accepted naturalism. According to Zahar, Popper neither contradicted himself nor changed his mind, but instead remained consistent over time.

Zahar claims that when Popper seems to argue that methodology is a science like any other which can be tested against the facts of scientific behaviour, the facts he has in mind are "normative facts" (Zahar's term). Normative facts are facts about how science is best conducted, and are thus different from empirical facts about how science is actually conducted. Popper's method is "quasi-empirical" (Popper's term) because it uses "normative facts" to refute methodologies in a manner analogous to his falsificationist account of the refutation of scientific theories. If a methodology says that in a particular situation scientists should act in a certain way, and this conflicts with a normative fact, then the methodology must be rejected as unsatisfactory. It seems reasonable to assume that if, on the other hand, a methodology implies that scientists ought to behave in a way that they actually do, over a wide range of circumstances, then its claim to acceptance is strengthened, in analogy to Popper's corroboration criterion.

According to Zahar's interpretation then, Popper's method of assessing methodologies is analogous to his methodology, but it is not the same. When Popper says that we should compare our methodologies with scientific practice he is not referring to just any scientific practice, but good scientific practice. If Popper does in fact make this distinction, then we can see that he does not contradict himself with respect to the empirical status of methodology:

> We are now entitled to ask: is Popper's position, taken over time, coherent? Did he not adhere in 1930-1 to the same naturalism which he later rejected in the *Logic of Scientific Discovery* only to adopt it implicitly in *Conjectures and*

> *Refutations*? I think not. According to Popper, methodologies should be based not on empirical, but as it were on normative, facts. On the one hand, Popper takes account of science as pursued, not by anybody, but by a carefully selected group of people. Not only would Popper not want to treat Newton and Nostradamus on a par, he would want to include the one and explicitly exclude the other... On the other hand, Popper expects his methodology to imply not *factual*, but normative, propositions like: Einstein's hypotheses are more 'scientific' than Freud's; a corroborated theory is preferable to a refuted one; general relativity is more strongly supported by the star shift than by freely falling bodies; modern physics constitutes progress over classical mechanics *etc.* (Zahar 1983, p. 153)

We are now in a position to understand what is meant by the "quasi-empirical" aspect of Popper's method. It is not purely empirical, because the statements used to refute and corroborate methodologies are normative (for example, theory A is better than theory B). Thus Popper's appeal to the practice of science in defence of his methodology does not mean that he is adopting a naturalistic position because he is referring to "good" scientific practice, which is not an empirical concept. Popper is not referring to a natural fact about how science *is* conducted but a non-empirical account of how science *should* be conducted.

This procedure is analogous to one in moral philosophy. We might say that the general principle "Thou shalt not kill" is refuted by the statement "It is right to kill Hitler". The universal rule "Thou shalt not kill" implies that "It is wrong to kill Hitler." If it is right to kill Hitler, then we may consider "Thou shalt not kill" refuted. Neither is an empirical statement: they are instead normative statements (assuming, of course, a non-naturalist approach to ethics, which at least seems coherent). The refutation of the principle "Thou shalt not kill" by the particular "It is right to kill Hitler" is not a matter of empirical consideration, but of normative considerations, although it is like an empirical refutation insofar as it uses modus tollens.

2. 'Normative facts'

To complete Zahar's account of Popper we must consider how we come to know particular "normative facts," that is, what sorts of reasons might be cited for the judgment that, say, Newtonian mechanics is better than Nostradamus's oracular predictions. As we have seen, Popper's answer is that the definition of "good science" is conventional, so that such value judgements are conventional too. Given initial agreement on what good science is, methodologies can be assessed in light of how well they capture this science. But Popper, as Zahar points out, holds that initial agreement cannot be provided by rational argument:

> Such value-judgments are not *directly available* in the same way that basic statements are given, but are *conferred* by the mind on certain achievements. Unlike physical theories, methodological systems are comparable only if they

> share certain normative propositions. It is impossible to refute a methodologist who insists that science ought to be restricted to what is directly given by sense-experience even at the cost of regarding classical physics, relativity theory and quantum mechanics as illegitimate speculations. (Zahar 1983, p. 153)

Zahar cites a passage from *Die Beiden Grundprobleme* in support of this interpretation:

> What does scientific success consist in? We are far from believing that a theoretical answer can be given to this question. We rather say that the answer depends on what is regarded as scientifically valuable. Methodology [*Methoden-Theorie*] will thus become a discipline which starts from certain scientific values, or to use a sober formulation, from certain scientific aims and ends. There can be many very different such aims and ends. I consider a rational decision between them impossible. One can set as an aim for science the construction of the safest possible theories, even maybe of an absolutely secure theory. To a person who pursues such an aim the development of physics since the turn of the century must appear as a breakdown of science [Popper goes on to say that Dingler was such a person]....If we nonetheless plead for modern natural science, it is not because science factually is what it is, but, to put it crudely, because we like science as it is. We simply appraise differently from Dingler. (Popper, Zahar's translation, Zahar 1983, p. 153-154)

If we agree on what good science is, then we can conduct the business of finding the methodology that captures this. If not, there can be no further (rational) debate.

However, if this interpretation is correct, Zahar's "normative facts" are not facts as normally understood. They are not adopted by observing external reality, but through a conventional choice based on personal preference, not rational argument. It seems that the "truth values" of such "facts" can be determined only by reference to our personal preferences. This is analogous to relativism or emotivism in ethics, where moral judgements are taken to be an expression of an individual's preferences, not as descriptive of any moral reality. It follows that if these "facts" have truth values, then methodology becomes an empirical theory about our preferences. This is surely not what Popper (or Zahar) had in mind.

The reference to "normative facts" seems to contradict Popper's views of what methodologies are based on. Popper's argument against naturalism is that it treats a conventional decision as to what good science is as if it were a fact, and that this leads to dogmatism. Popper clearly holds that methodologies are based on conventions which do not have a truth value. I conclude that we should understand what Zahar terms "normative facts" not as facts, but as conventions. It does seem that, given this interpretation, Popper is consistent. He only refers to scientific episodes because he deems them "good", not because they lend empirical support to his methodology.

This leads to the problem I referred to earlier, namely, that with this interpretation

Popper can give offer no justification for choosing one method over another. If the choice of a methodology is based on personal preference, and if there is disagreement as to these, then there seems to be no reason for choosing one methodology over another.

So, Popper's alternative to a naturalistic approach which judges methodology by applying scientific standards is a conventional methodology. I will postpone the question of whether or not such a methodology is adequate at all until the final chapter. For now, however, we may conclude that Popper is a consistent non-naturalist, if interpreted as a conventionalist.

F. Methodological statements as basic statements

Before I continue, I must consider one possible way of adjudicating between methodologies which would be consistent with Popper's conventionalism (which Popper himself does not consider). Popper could argue that methodologies may be tested by statements analogous to his "basic statements." This seems a reasonable strategy because Popper often draws analogies between his methodology and his meta-methodology.

Popper's methodology requires that scientific theories be falsifiable. This raises the question of the form of the falsifying statement. According to Popper, the falsifying statement must be "singular": "If falsifiability is to be at all applicable as a criterion of demarcation, then singular statements must be available which can serve as premises in falsifying inferences." (1968, p. 43) But to apply modus tollens and thus arrive at the conclusion that a certain theory is false, we must know that the singular statements used to falsify are true. Popper denies that we can know for sure that any such proposition is true, for this would involve induction. Thus it would seem that we can never falsify a theory using his method. Popper's response is that we make a conventional choice to treat certain singular statements as if they were true. Popper terms these "basic statements": "Every test of a theory, whether resulting in its corroboration or falsification, must stop at some basic statement or other which we *decide to accept*. If we do not come to any decision, and do not accept some basic statement or other, then the test will have led nowhere." (Popper 1968, p. 104) Popper acknowledges that the choice of a basic statement is conventional: "Basic statements are accepted as the result of a decision or agreement; and to that extent they are conventions." (Popper 1965, p. 106):

> But considered from a logical point of view, the situation is never such that it compels us to stop at this particular basic statement rather than at that, or else give up the test altogether. For any basic statement can again in its turn be subjected to tests, using as a touchstone any of the basic statements which can be deduced from it with the help of some theory, either the one under test, or another. This procedure has no natural end. Thus if the test is to lead us any-

> where, nothing remains but to stop at some point or other and say that we are satisfied, for the time being. (Popper 1968, p. 104)

Popper calls basic statements "dogmas": "The basic statements at which we stop, which we decide to accept as satisfactory, and as sufficiently tested, have admittedly the character of *dogmas*, but only in so far as we may desist from justifying them by further arguments (or by further tests)." (Popper 1965, p. 105) But, he claims, this "kind of dogmatism is innocuous since, should the need arise, these statements can easily be tested further." (1965, p. 105)

It might seem possible to construct a defence of Popper's methodology from the charge of relativism using basic statements as an analogy. It could be argued that science as it conventionally defined is like a basic statement which is conventionally accepted. We agree on a definition, but we can, "should the need arise", revise it. The initial agreement concerning the definition of "good science" would be conventional, but the continued acceptance of a particular definition of science would be determined by rational argument, and so a choice of methodology need not depend solely on personal preference.

But there is a significant disanalogy between basic statements and methodological statements. It is clear how we could further test the basic statements of Popper's methodology. Suppose that we wished to test some general theory about electricity, and that this theory seems to be falsified because the meter does not give the predicted value. Then we may make the conventional choice to decide that our meter reading is in fact correct. The basic statement would presumably be of the sort "such and such amount of electricity is flowing through the circuit." However, we could test this, presumably by conducting experiments on meters of the appropriate kind.

No analogous procedure, it seems, is available for reevaluating methodological conventions. The only statement that could contradict a methodology would be, according to Popper, another conventional statement. The only reason that seems available to Popper if he is to be true to his conventionalism is that we do or do not like a particular definition of science. There is no other general theory to which we can appeal other than our preferences.

G. Popper's naturalism

Popper seems to wish to avoid the implication of the conventionalist approach that methdologies are adopted on the basis of personal preferences. There are, I have argued, two alternatives to the conventionalist approach: apriorism and aposteriorism. The apriori approach to methodology faces severe problems, as Popper acknowledges. So it seems that Popper must appeal to empirical considerations in support of his methodology. In this section I will argue that this is in fact what Popper does, and that his methodology must rely on empirical evidence for support. In *Die Beiden Grundprobleme* Popper proposes an "evolutionary defence" of his methodology. This defence seems to be naturalistic, because it appeals to empirical facts about natural selection:

> We can defend our appraisal *in no other way* but in saying that it corresponds to our world-view, that it corresponds to the biological role played by science in this world-view: science is the engineers corps of natural adaptation which is sent furthest forward; this is why it must be exposed to natural selection. *That modern science agrees to a large extent with our methodological ideal* can be naturally [*zwanglos*] explained in our world view as the outcome of natural selection. (Popper, Zahar's translation 1983, p. 154, my italics)

This clear use of empirical facts in support of his methodology would be compatible with his conventionalism if Popper viewed the theory of natural selection as good by convention, as it were. If we have accepted a particular theory of natural selection, then, according to Popper, his methodology is in accord with this theory. However, in *The Logic of Scientific Discovery* Popper tells us that methodologies may be judged by the degree to which they conform to scientists' intuitions:

> *My only reason* for proposing my criterion of demarcation is that it is fruitful: that a great many points can be clarified and explained with its help. 'Definitions are dogmas; only the conclusions drawn from them can afford us any new insight', says Menger. This is certainly true of the definition of the concept 'science'. It is only from the methodological decisions which depend upon this definition, that the scientist will be able to see how far it conforms to his intuitive idea of the goal of his endeavours. (Popper 1965, p. 55, my italics)

And Popper claims that science uses the same methods as the philosopher to assess methodologies:

> The philosopher too will accept my definition as useful only if he can accept its consequences. We must satisfy him that these consequences enable us to detect inconsistencies and inadequacies in older theories of knowledge, and to trace these back to the fundamental assumptions and conventions from which they spring. But we must also satisfy him that our own proposals are not threatened by the same kind of difficulties. *This method of detecting and resolving contradictions is applied also within science itself, but it is of particular importance in the theory of knowledge. It is by this method, if by any, that methodological conventions might be justified, and prove their value.* (Popper 1965, p. 55, my italics)

It is clear that Popper's defence of his methodology is caught between the two contradictory positions: the view that methodologies are conventional and the view that they are empirical. He first claims that a choice of methodology cannot be rationally debated, thus excluding an apriori choice. However, he then appeals to empirical support for justification (by appealing to natural selection and the intuitions of scientists).

H. Conclusion

Popper rejects both apriori and naturalistic approaches to methodology in favour of conventionalism. But this has the difficulty that the choice among methodologies is not rational, but a matter analogous to taste. It is this unattractive implication that seems to have propelled Popper back to a naturalistic approach and to cite scientific practice in support of his methodology.

Popper argues that a naturalistic approach is untenable because methodological statements are conventional. But, Popper at times seems to hold to the naturalistic view that empirical evidence should influence a choice of methodology. This contradiction emphasises the appeal of an approach which does account for how scientists actually act. However, I shall argue that we can have reasons for choosing between conventions. In the next and final chapter I shall put forward some tenative suggestions along this line.

CHAPTER 8: Conclusion

In this thesis I have considered three approaches to methodology—apriori, aposteriori and conventional. I settle on conventionalism for the following reason: as I argued in Parts I and II, neither the apriori nor the aposteriori approach is workable. At the end of Part I, I argued that the apriori approach cannot succeed, since there will always be the possibility of disagreement. In Part II I argued that the aposteriori approach is caught in a dilemma: it is either self-refuting or argued for in a circular manner. My argument for conventionalism is thus negative: no other approach can succeed. It might appear that I have not defined conventionalism exhaustively, and that consequently I do not establish that conventionalism is indeed a genuine alternative to the other approaches. This is incorrect: I have presented three conceptions of conventionalism, as represented by the views of Carnap, Quine and Popper, who have, at times, taken a conventionalist line. I have not defended one particular conventionalist approach: to do so would require another thesis. Nonetheless, the writings of the three above mentioned authors indicate how conventionalism can be applied. In fact, they show that, for the most part, conventionalist methodology is not be radically different from other types of conventionalism.

There is no one conventionalist method: Carnap advocated what can be called an apriori-cum-conventionalist approach, while Quine and Popper advocated what can be called an empirical-cum-conventionalist line. Carnap espoused a conventionalist line, but regarded logical considerations as foremost in methodology. Similarly Quine considers empirical considerations to be of the highest importance, and Popper held them to be less important. The different approaches illustrate a basic difficulty for conventionalism: the question as to what counts as a reason for a methodology. A step towards answering this question might perhaps be found in a careful study of, for example, debates in probabilistic methodology. Quine based his theory of methodology on a behaviourist theory of learning. As I argued in Chapter 6, this is the same as a behaviourist interpretation of the Dutch Book argument. But in Chapter 3 I argued that adopting such an interpretation makes the argument invalid. This debate is based on assumptions concerning dispositions to behaviour which are, at present, perhaps uncontroversial. But, these assumptions may be further debated. In this way, methodological statements do seem analogous to Popper's basic statements, in that they are conventionally accepted as true. This still leaves open the question of what is to count as a reason. In science, empirical evidence gives the grounds for choosing a particular theory. In methodology, we seem to have no such evidence. This is incorrect however: it is just the choice of what counts as evidence is based on conventions. Once these are accepted, we can adjudicate between methodologies.

Indeed, this is how much methodology is in fact carried out. The only difference is that a conventionalist explicitly acknowledges that there may be no end to the debate. It thus seems that we can have reasons for choosing between methodologies, although these reasons are pragmatic, in the sense that they are based on scientific concerns. The fundamental question of epistemology seems to be: 'What is the best way to learn from experience?' There can be no final anwer to this question, as I argued in Part I. Nor can the question be entirely answered in a scientific fashion, as I argued in Part II. Instead, the question can at least be addressed if we begin from some fundamental, agreed upon starting points, which can in turn always be debated. Some might object to this conclusion, and argue that it leaves methodology without foundations. My response is that if the arguments of Parts I and II are correct, we have no other choice.

Bibliography

Almeder, Robert, 1990, "On Naturalizing Epistemology" *American Philosophical Quarterly* **27** (4), pp. 263-279.

Anand, Paul, 1993, *Foundations of rational choice under risk* (Oxford: Oxford University Press).

Bateman, Bradley W., 1988, "G.E. Moore and J.M. Keynes: a missing chapter in the history of the expected utility model", *American Economic Review*, **78** (5), pp. 1098-1106.

Berlin, I., 1939, *Karl Marx: His Life and Environment* (London: Thornton Butterworth Ltd.)

Borel, Émile, 1965, *Elements of the theory of probability,* trans. John E. Freund, 1950 (Englewood Cliffs: Prentice-Hall).

Braithwaite, R.B., *Scientific Explanation: A study of the function of theory, probability and law in science* (Cambridge: Cambridge University Press), 1968.

Carnap, Rudolf, 1928, *The Logical Structure of the World and Pseudoproblems in Philosophy*, trans. R. George, repr. 1967 (Berkeley: University of California Press).

Carnap, Rudolf, 1932, "The Elimination of Metaphysics through the Logical Analysis of Language", trans. Arthur Pap, repr. A. Ayer, ed., 1959, *Logical Positivism* (New York: The Free Press), pp. 60-81.

Carnap, Rudolf, 1934, "On the Character of Philosophic Problems" *Philosophy of Science*, vol. 1, pp. 5-19, trans. W.M. Malisof, repr. in Rorty 1967, pp. 54-62.

Carnap, Rudolf, 1936, "Testability and Meaning", *Philosophy of Science* **3** (4), pp. 419 - 71.

Carnap, Rudolf, 1945, "On Inductive Logic", *Philosophy of Science* **12**, pp. 72-97.

Carnap, Rudolf, 1950, "Empiricism, Semantics, and Ontology" *Revue Internationale de Philosophie* **4**, pp. 20-40, repr. in Rorty 1967, pp. 72-84.

Carnap, Rudolf, 1950, *Logical Foundations of Probability* (Chicago: University of Chicago Press).

Carnap, Rudolf, 1952, *The Continuum of Inductive Methods* (Chicago: University of Chicago Press).

Carnap, Rudolf, 1962, "The Aim of Inductive Logic", *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress,* eds. Ernest Nagel, Patrick Suppes and Alfred Tarski (Stanford: Stanford University Press), pp. 303-318.

Carnap, Rudolf, 1963, *The Philosophy of Rudolf Carnap*, ed. P. Schilpp (LaSalle: Open Court).

Carnap, Rudolf, 1968, "Inductive Logic and Inductive Intuition" *The Problem of Inductive Logic*, ed. I. Lakatos (Amsterdam: North-Holland Publishing Co.) pp. 258-267.

Carnap, Rudolf, 1971, "Inductive Logic and Rational Decisions", *Studies in Inductive Logic and Probability*, vol.1, R. Carnap and R. Jeffrey, eds., (Berkeley: University of California Press), pp. 5-31.

Carabelli, A.M., 1988, *On Keynes's Method* (Basingstoke: Macmillan).

Chomsky, Noam, 1975, "Knowledge of Language," *Language, Mind, and Knowledge: Minnesota Studies in the Philosophy of Science vol VII*, ed. Keith Gunderson (Minneapolis: University of Minnesota Press), pp. 299-320.

Cosmides, Leda and Tooby, John, Manuscript, "Are humans good statisticians after all?: Rethinking some conclusions from the literature on judgement under uncertainty."

and ed.
Creath, Richard, 1990, "Introduction," *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work* (Berkeley: University of California Press).

Dennett, Daniel C., 1992, *Consciousness explained* (London: Allen Lane).

de Finetti, B., 1974, *Theory of probability : a critical introductory treatment*, trans. Antonio Machi and Adrian Smith (London : Wiley).

Feldman, Fred, 1978, *Introductory Ethics* (Englewood Cliffs, N.J.: Prentice-Hall, Inc.).

Frankena, William K., 1963, *Ethics* (Englewood Cliffs, N.J.: Prentice-Hall, Inc.).

Giere, Ronald N., 1985, "Philosophy of Science Naturalized", *Philosophy of Science,* **56**, pp. 331-356.

Giere, Ronald N., 1988, *Explaining Science: a cognitive approach* (Chicago: University of Chicago Press).

Giere, Ronald N., 1989, "Scientific Rationality As Instrumental Rationality", *Studies in the History and Philosophy of Science*, **20** (3), pp.377-384.

Gigerenzer, G., 1991, "How to make cognitive illusions disappear: Beyond heuristics and biases, *European Review of Social Psychology* **2**, pp. 83-114.

Gillies, Donald Angus, 1973, *An Objective Theory of Probability* (London: Methuen).

Gillies, Donald Angus, 1988, "Induction and Probability" *An encyclopaedia of philosophy*, ed. G.H.R. Parkinson (London: Routledge), pp. 179-204.

Gregory, Richard L., 1987, "Yellow" *The Oxford Companion to the Mind*, ed. Gregory, Richard L., (Oxford: Oxford University Press)

Herrnstein, R.J., 1982, "Stimuli and the texture of experience" *Neuroscience and Biobehavioral Reviews* 6 (1), pp. 105-117.

Howson, Colin, 1976, "The Development of Logical Probability", *Essays in Memory of Imre Lakatos*, ed. R.S. Cohen, P.K. Feyerabend and M.W. Wartofsky, (Dordrecht: D. Reidel Publishing Co.), pp. 276-298

Howson, Colin and Peter Urbach, 1993, *Scientific Reasoning: The Bayesian Approach* 2nd ed. (Chicago: Open Court).

Kemeny, John G, 1955, "Fair Bets and Inductive Probabilities" *The Journal of Symbolic Logic* **20** (3), pp. 263-273.

Keynes, John Maynard, 1921, *A Treatise on Probability,* repr. 1973, *The collected writings of John Maynard Keynes*, vol. 8 (London: Macmillan for the Royal Economic Society).

Keynes, John Maynard, 1933, *Essays in Biography* [2nd ed.], repr. 1972, *The collected writings of John Maynard Keynes*, vol. 10 (London: Macmillan for the Royal Economic Society).

Kim, Jaegwon, 1988, "'What is "Naturalized Epistemology?'", *Philosophical Perspectives: Epistemology* **2**, pp. 381-405.

Kornblith, Hilary, 1985, "What is Naturalistic Epistemology?" *Naturalizing epistemology*, ed. Hilary Kornblith (Cambridge, Mass.: MIT Press), pp. 1-13.

Kuhn, Thomas, 1970, "Logic of Discovery or Psychology of Research?" and "Reflections on My Critics" in Lakatos and Musgrave 1970, pp. 1-23, pp. 231-278.

Lakatos, Imre, 1970, "Falsification and the Methodology of Scientific Research Progammes" in Lakatos and Musgrave 1970, pp. 91-196.

Lakatos, Imre and Alan Musgrave, 1970, eds., *Criticism and the Growth of*

Knowledge (Cambridge: Cambridge University Press).

Laudan, Larry, 1987, "Progress or Rationality? The Prospects for Normative Naturalism" *American Philosophical Quarterly* **24** (1), pp. 19-31.

Lehman, R. Sherman, 1955, "On Confirmation and Rational Betting" *The Journal of Symbolic Logic* **20** (3), pp. 251-262.

Levy, Paul, 1975, "The Bloomsbury Group" *Essays on John Maynard Keynes*, ed. Milo Keynes (Cambridge: Cambridge University Press), pp. 60-72.

MacIntyre, Alasdair, 1984, *After Virtue*, 2nd ed. (Notre Dame, Indiana: Notre Dame Press).

Maffie, James, 1993, "Recent Work on Naturalized Epistemology" *American Philosophical Quarterly*, **27** (4), pp. 281-293.

Mises, Richard von, 1957, *Probability, Statistics and Truth*, 2nd rev. English ed., prepared by Hilda Geiringer, repr. 1981 (New York: Dover Publications, Inc.)

Moggridge, D. E., 1992, *John Maynard Keynes: an economist's biography* (London: Routledge).

Moore, George Edward, 1903, *Principia Ethica*, repr. 1968 (Cambridge: Cambridge University Press).

Neyman, J., 1952, *Lectures and Conferences on Mathematical Statistics and Probability*, 2nd ed. (Washington: U.S. Department of Agriculture).

Nozick, Robert, 1981, *Philosophical Explanations* (Oxford: Clarendon Press).

O'Donnell, Rod M., 1989, *Keynes: philosophy, economics and politics: the philosophical foundations of Keynes's thought and their influence on his economics and politics* (Basingstoke: Macmillan).

O'Donnell, Rod M., 1991, "Keynes on Probability, Expectation and Uncertainty" *Keynes as Philosopher-Economist: The Ninth Keynes Seminar held at the University of Kent 1989*, ed. Rod M. O'Donnell (London: MacMillan), pp. 3-60.

Popper, Karl R., 1965, *Conjectures and refutations: the growth of scientific knowledge* (New York: Harper and Row).

Popper, Karl R., 1968, *The logic of scientific discovery* 2nd revised ed. 193~~2~~4 *Logik der Forschung*, (New York: Harper and Row).

Popper, Karl, 1970, "Normal Science and its Dangers" in Lakatos and Musgrave 1970, pp. 51-58.

Popper, Karl R., 1983, *Realism and the Aim of Science: from the Postscript to The Logic of Scientific Discovery*, ed. W.W. Bartley III (London: Routledge)

Putnam, Hilary, 1982, "Why Reason Can't be Naturalized" *Philosophy, Mind and Cognitive Inquiry*, eds. David J. Cole, James H. Fetzer and Terry L. Rankin (Dordrecht: Kluwer Academic Publishers), pp.283-303.

Quine, W.V.O., 1953, "Two Dogmas of Empiricism" *From a Logical Point of View*, 2nd revised edition 1980 (Cambridge, Mass: Harvard University Press), pp. 20 - 46.

Quine, W.V.O., 1953a, "On What There Is" *From a Logical Point of View*, 2nd revised edition 1980 (Cambridge, Mass: Harvard University Press), pp. 1 - 19.

Quine, W.V.O., 1960, *Word and Object* (Cambridge, Massachusetts: The Technology Press of the Massachusetts Institute of Technology).

Quine, W.V.O., 1966 *The Ways of Paradox, and other essays*, revised and enlarged ed. 1976 (Cambridge, Mass.: Harvard University Press)

Quine, W.V.O., 1969, "Epistemology Naturalized", *Ontological Relativity and Other Essays* (New York: Columbia University Press), pp. 69-90.

Quine, W.V.O., 1969a, "Natural Kinds", *Essays in Honor of Carl G. Hempel: A Tribute on the Occasion of his Sixty-Fifth Birthday*, ed. Nicholas Rescher (Dordrecht: D. Reidel Publishing Co.), pp. 5-23.

Quine, W.V.O., 1970, "On the reasons for indeterminacy of translations" *Journal of Philosophy* **67**, pp. 178 - 183.

Quine, W.V.O., 1970a, *Philosophy of Logic* (Englewood Cliffs, New Jersey: Prentice Hall).

Quine, W.V.O., 1972, "Methodological Reflections on Current Linguistic Theory", *Semantics of Natural Language*, eds. D. Davidson and G. Harman (Dordrecht: D. Reidel Publishing Co.), pp. 442-454.

Quine, V.V.O., 1973, *The Roots of Reference* (La Salle, Illinois: Open Court).

Quine, W.V.O., 1975, "The Nature of Natural Knowledge", *Mind and Language: Wolfson College Lectures 1974*, ed. Samuel Guttenplan (Oxford: Clarendon Press), pp. 67-81.

Quine, W.V.O., 1978, "Facts of the Matter"*American Philosophy from Edwards to Quine*, ed. R. Shahan (Norman, Oklahoma: University of Oklahoma Press), pp. 176-196, repr. 1979 *The Southwestern Journal of Philosophy*, **9** (2), pp. 155-169.

Quine, W.V.O., 1981 "Reply to Stroud", *Midwest Studies in Philosophy*, eds. P.A. French, T.E. Uehling and A.K. Wettstein (Minneapolis: University of Minnesota Press), pp. 473-475.

Quine, W.V.O., 1986, *The Philosophy of W.V.O. Quine*, eds. Lewis Edwin Hahn and Paul Arthur Schilpp (La Salle, Illinois: Open Court).

Quine, W.V.O., 1987, *Quiddities...*

Quine, W.V.O., 1992, *Pursuit of Truth* (Cambridge, Massachusetts: Harvard University Press).

Quine, W.V.O., 1993, "In Praise of Observation Sentences" *The Journal of Philosophy*, **90** (3), pp.107-116.

Quine, W.V.O., 1994, "Promoting Extensionality" *Synthese* **98**, pp. 143-151.

Quine, W.V.O. and Ullian, J.S., 1970, *The Web of Belief* (New York: Random House).

Ramsey, Frank Plumpton, 1978, *Foundations: essays in philosophy, logic, mathematics and economics*, ed. D.H. Mellor (London: Routledge & Kegan Paul).

Rorty, Richard, 1967, ed. *The Linguistic Turn: Recent Essays in Philosophical Method* (Chicago: The University of Chicago Press).

Rorty, Richard, 1980, *Philosophy and the mirror of nature* (Oxford: Blackwell).

Russell, Bertrand, 1910, "Knowledge by Acquaintance and Knowledge by Description" *Philosophical Essays* repr. 1917 *Mysticism and Logic* (London: George Allen and Unwin Ltd.).

Russell, Bertrand, 1912, *The Problems of Philosophy* (London: Oxford University Press) reset in 1946.

Russell, Bertrand, 1967, *The autobiography of Bertrand Russell, vol. 1: 1872-1914* (London: Allen and Unwin).

Sarkar, Husain, 1983, *A Theory of Method* (Berkeley : University of California Press).

Savage, Leonard J., 1972, *The Foundations of Statistics* 2nd. revised ed. (New York: Dover Publications).

Schick, Frederic, 1986, "Dutch Bookies and Money Pumps" *Journal of Philosophy*, 83, pp.112-119.

Siegel, Harvey, 1980, "Justification, Discovery and the Naturalizing of Epistemology", *Philosophy of Science* 47, pp.297-321.

Siegel, Harvey, 1989, "Philosophy of Science Naturalized? Some Problems with Giere's Naturalism" *Studies in the History and Philosophy of Science* 20 (3), pp.365-375.

Shimony, Abner, 1955, "Coherence and the Axioms of Confirmation" *The Journal of Symbolic Logic*, 20 (1), pp. 1-28.

Skidelsky, Robert, 1983, *John Maynard Keynes: Hopes betrayed 1883-1920*, vol.1 (London: Macmillan).

Skyrms, Brian, 1980, *Causal necessity: a pragmatic investigation of the necessity of laws* (New Haven: Yale University Press).

Skyrms, Brian, 1986 *Choice and chance: an introduction to inductive logic,* 3rd ed., (Belmont, California: Wadsworth Pubublishing Co.).

Stroud, Barry, 1984, *The Significance of Philosophical Scepticism* (Oxford: Clarendon).

Székely, Gábor J., 1986, *Paradoxes in probability theory and mathematical statistics* (Dordrecht: Reidel).

Urbach, Peter, 1988, "What is a Law of Nature? A Humean Answer" *British Journal for the Philosophy of Science* 39, pp.193-210.

Warnock, G.J., 1967, *Contemporary Moral Philosophy* (London: Macmillan).

Zahar, Elie, 1983, "The Popper-Lakatos Controversy in the light of *Die Beiden Grundprobleme der Erkenntnistheorie*" *British Journal for the Philosophy of Science* **34**, pp. 149-171.

ABSTRACT

The relation between the normative and the empirical in the philosophy of science is examined by investigating apriori and aposteriori approaches to methodology. The apriori is usually equated with the prescriptive, and the aposteriori with the descriptive. It is argued that this equation is mistaken, and that neither a purely apriori nor a purely aposteriori approach to methodology can succeed. Methodologies based on probability are used as illustrations.

Purely apriori and purely aposteriori approaches are examined in Parts I and II respectively. The former are investigated through the intuitionism of J.M. Keynes and the analytic method of Carnap. Dutch Book arguments are also considered as apriori arguments. I conclude that an apriori approach is irredeemably flawed, in that it can never meet the goal it sets for itself of producing a self-evidently justified set of rules for science. Purely aposteriori approaches are investigated in the second Part by focussing on R. Giere's and W.V.O. Quine's proposals for a naturalised epistemology. It is argued that a purely empirical approach is caught on the horns of a dilemma: if it is defended on aposteriori grounds then the argument is circular, and if on apriori grounds it is self-refuting. Thus it is shown that the aposteriori approach too cannot serve as the foundations for methodology.

However, I shall argue that Quine's project has been misunderstood, and that in fact Quine argues for aposteriori methodology from conventionalist grounds. The possibility of a conventionalist approach to the philosophy of science which avoids the problems of the purely empirical and of the purely apriori approach is explored in the third Part of this thesis. Karl Popper's early advocacy of such a conventionalist approach is discussed. The final chapter is devoted to showing how certain flaws in Popper's and Quine's conventionalist approaches may be mended. It is concluded that the conventionalist approach to methodology provides an adequate framework for the relation between the normative and the empirical in the philosophy of science.

formulating an acceptable conventionalism, so I shall examine two proposals put forward by Quine and by K.R. Popper. The last chapter of this thesis is a tentative suggestion along lines suggested by these authors.

Such a grand project, that of showing that only conventionalism has so far survived, and that it is the only view that can survive, is obviously impossible (even neglecting the fact that it is undertaken by a single doctoral candidate). I limit my project in two ways. First, I concentrate only on what might be called empiricist epistemology by restricting the scope of the thesis to methodology. Secondly, I for the most part consider only probabilistic methodologies. I feel that much will not be overlooked given these constraints since most of what I say about probabilities applies *mutatis mutandis* about other methodologies.

Specifically: Part 1 deals with the issues around methodology considered as an apriori discipline. Chapter 1 covers intuitionism, with the main protagonists being J.M. Keynes and G.E. Moore. Chapter 2 deals with the analytic method proposed by Rudolf Carnap. I argue that both attempts are failures. In Chapter 3 I argue that the further failures in this programme show apriorism to be an unsuccessful enterprise.

Part II deals with the purely empirical approach. I examine and dismiss in Chapter 4 what may be called 'naive' empiricism (or 'naturalism,' as empirical methodologies are sometimes called). Only one person, R. Giere, really seems to have ever held this view: W.V.O. Quine, who is often cited as a proponent of this view instead holds, I shall argue in Chapter 5, like Carnap, an empirical-cum-conventionalist view. While I shall later argue that this view is broadly correct, I conclude that Quine's version is too narrow conceived.

The shortest section of this thesis is Part III, which is an attempt to show how to work towards an acceptable conventionalism. Chapter 6 is an argument against Karl Popper's apriorism-cum-conventionalism. Chapter 7 is an attempt to show that conventionalism is an acceptable alternative.

Acknowledgements

I cannot imagine that there could be a finer tutor than Peter Urbach: even if I could list his many qualities, they would be too many to include here. His detailed and careful criticisms of this thesis were invaluable, as is his thoughtful advice and steadfast support.

Colin Howson and Thomas Uebel have both discussed many of the issues in this thesis with me. I appreciate their advice and encouragement. Professor Nancy Cartwright discussed with me many issues related to those found in this thesis, and helped me obtain both financial support and work.

I wish I knew how to thank properly my wife, Hope Marie Childers, and my parents. I am grateful for their support and forbearance of the vicissitudes of academic life, and for their love.

I am grateful to the members of the Philosophy of Science Research Students Group for organising the best series of seminars I have ever attended. I am particularly grateful to Marco del Seta and Samet Bagce for their constant willingness to engage in ferocious argument in convivial surroundings.

Husain Sarkar first introduced me to the issues with which this thesis is concerned, and encouraged me to study at the LSE. Ms. Theresa Hunt and Mrs. Pat Gardner helped in negotiating necessary bureaucratic hurdles. The All-London Centre for the Philosophy of the Natural and Social Sciences provided me with a desk. Many friends in London have supported me in difficult times: Imogen Planner, Austen Garth, Robin Hendry, Dagmar Lorenz-Meyer, Joanna Bennet and Graham Combi in particular. I thank them all.

Finally, I thank the members of the Working Group in Logic, Institute of Philosophy, Academy of Sciences of the Czech Republic: particular thanks are due Vladimír Svoboda, Ondrej Majer and Petr Kolář, for showing me my new home.