COMPARISON AND COMPETITION IN SCIENTIFIC THEORIES

SUBMITTED BY

JOHN NELSON PERRY, B.Sc.,

FOR THE DEGREE OF

MASTER OF LETTERS

FROM THE UNIVERSITY OF ST. ANDREWS.

St. Andrews, June 1972.

Acknowledgements

I would like to thank Dr. Eva Cassirer, whose helpful criticism during the course of this work was only matched by her untiring patience. The many errors remaining are, of course, entirely my responsibility.

I also thank my wife, Lalage,without whose moral and physical support this work would probably never have been completed.

JNP.

TABLE OF CONTENTS.

Scientific theories very rarely die from neglect. They are removed from our body of conjectural knowledge by abscission. New theories compete with the old for acceptance and, if they are successful in gaining the scientist's approval, lead to the rejection of the old. This essay is an attempt to describe the principles of selection which govern this struggle for survival.

After a discussion of terms associated with competitions I shall try to differentiate the various methods of comparison which are applicable to theory assessment. In Chapter Two I shall examine why some instances of these methods are unsatisfactory because of problems associated with language dependancy. In Chapter Three I shall, while discussing comparisons based on 'falsificationist' priciples, try to make good part of my claim that any philosophically adequate account of the growth of knowledge must include instances of each of the three methods of comparison which I describe. In Chapter Four I shall consider the influence of Lakatos' 'sophisticated methodological falsificationism' on our assessment of other falsificationist comparisons.

To avoid a lengthy aside in the middle of this essay, I shall begin with a brief summary of the 'sophisticated methodological falsificationist' position put forward by Lakatos.[1]

The eighteenth century empiricists used the term 'knowledge' to refer both to factual propositions and to true theories which could, they believed, be infallibly proved from them. The precise nature of the mode of inference from factual propositions to true theories was never exactly formulated. It appears that they did not differentiate between deductive and inductive inference. It was agreed that any non-analytic proposition which had not had the stamp of infallibility confered upon it, either directly 'by experience', or indirectly by inference from factual propositions, did not constitute knowledge but was merely sophistry and illusion.

This led directly to the problem of the status of theoretical propositions. For the classical empiricist, propositions occuring within a theoretical framework were rigidly divided into two classes; factual or observational propositions and speculative or theoretical propositions. Observational propositions were held to be meaningful and, if proved true, constituted knowledge. Theoretical propositions relied for their claim to meaningfulness and knowledge, upon the possibility of reducing the theoretical terms occuring in them, to observational ones. Thus we see that the empiricist who wishes to establish the (absolute) truth of theories, relies upon two distinct inductive ascents. One from factual propositions to the universal generalisations of the theory, and another from factual or observational concepts to theoretical concepts.

Both these concepts of inductive ascent have been shaken irreparably. The former has suffered from the advance of deductive logic. It is now generally accepted that valid inference can only take the form of non-content increasing inference. Any inference which is content increasing (e.g. from a 'some' statement to an 'all' statement) is an informal inference and cannot guarantee the safe passage of the truth-value 'true' from the premiss to the conclusion. Thus the empiricist's inference from factual propositions to true theories, being content increasing, cannot confere the accolade of truth upon the theories so inferred.

The latter has not recovered from Duhem's criticism,[2] no less damning for it's being indirect. Implicit in the classical empiricist's account of the justification of the ascription of absolute truth values to theories, is his methodological principle regulating the growth of knowledge. We should start from indubitable factual propositions and self-evident observational concepts, and by gradual induction from these truths we arrive at higher truths. His methodological principle describes the accumulation of timeless knowledge. It follows that the meaning and the truth values of the theoretical concepts can never be inconsistent with the observational concepts upon which they are derivitive. What Duhem did was to demonstrate that just such an insonsistency did, in fact, occur and moreover in theories which were prized for their explanatory power. The introduction of novel theoretical terms led to inconsistency with the old observational laws which were based on concepts which were (according to the classical empiricist) used to derive the theoretical concept. Again we find that this empiricist account falls short of its intented aim.

These arguments resulted in a switch from classical epistemology to fallibilism. To preserve theoretical science as knowledge, the very concept of knowledge was radically altered. Theoretical science was held to be knowledge in a weaker sense, conjectural and hence fallible knowledge. Two problems arose, first the problem of how this conjectural knowledge was to be appraised and second, the problem of the growth of this knowledge. The attempt to answer the first of these problems led to the introduction of the idea that,,although no scientific theory can be proved to be true, they do have different probability. The ideal of this probabilistic philosophy is to assign to each theory, relative to the evidence available, a particular value of a probabilistic (in the sense of the calculus of probabilities) function which measures the degree of probability that each theory has, in the light of this evidence. I shall be discussing the role which such formal appraisals of theories can play in theory competition, in Chapter Two.

Probabilism does not have much to say about the second of the two problems mentioned; the problem of how conjectural knowledge can grow. The answer to this problem put forward by Popper does, however, have a great deal of relevance to the problem of appraisal. Popper approached his solution from a consideration of Hume's problem of induction. Hume originally formulated the logical problem; are we rationally justified in reasoning from repeated instances of which we have had experience to other instances of which we have had no experience? Hume answered 'no'. Moreover he claimed that it made no difference if, instead of asking for justification of certain belief, we merely asked for probable belief. Instances of which we have had experience do not allow us to reason or argue about the probability of non-experienced instances any more than to the certainty of such instances. But Hume also posed the psychological problem of induction. Why is it that all reasonable people believe that future instances will conform to the pattern they have experienced ? He answered this in terms of custom or habit, it was because of the irresistible power of the law of association.

Popper agreed with Hume's first answer but disagreed with the second. Because, he claimed, Hume's answers lead immediately to an irrationalist conclusion, he produced a reformulation of the problem which had happier results. Hume's irrational conclusion is that all our knowledge is just habit or custom, and is rationally totally indefensible. Popper's reformulation of the problem consists in an analysis of Hume's use of 'instance' and an extension of the problem of reasoning from instances to laws by taking into account counterinstances.

Regularities or laws are presupposed by Humes's term 'instance'; for, Popper claimed, the instances in question are instances <u>of</u> regularities or laws. Popper's reformulation of the logical problem of induction ammounts to: "Are we justified in reasoningfrom instances of which we have had experience, or from counterinstances of which we have had experience, to the truth or falsity of the corresponding laws and to future instances ? " Popper's answer illustrates the logical assymetry between confirmation and falsification. Although we are not justified in reasoning from an instance to the truth of the corresponding law, we <u>are</u> justified in reasoning from a counterinstance to the falsity of the corresponding law. If we accept a counterinstance, then logic forces us to accept that the corresponding law is false.

This result is the basis of Popper's methodological falsificationism and enabled him to establish a theory of method according to which it is possible to argue that one competing conjecture is preferable to another. Although it is true that every attempt to establish a general scientific law from instances is fallacious, this does not prevent us from considering one theory to be better than another. For example, although we cannot establish the truth of Newton's or Einstein's theories, this does not prevent us from saying that Einstein's is better, because there are counter-instances which refute Newton's theory but which do not refute Einstein's. By demonstrating, on purely rational grounds, how we can be justified in preferring some competing theories to others, Popper paves the way for a rational reconstruction of the growth of conjectural knowledge and an avoidance of Hume's scepticism.

This account of Popper's methodology is only a very crude outline. As I have portrayed it so far, it would seem that this brand of falsificationism suggests that science grows by the repeated overthrow of theories with the help of infallible information gleaned from facts. Whereas it admits the fallibility of all theories, it suggests that we can infallibily perform the act of repudiation of false theories. This is a position actually held by Medawar[3], but it is untenable. Firstly, it preserves the classical empiricist's rigid distinction between theoretical propositions of the one hand and factual or observational propoditions on the other. Popper denies the validity of this assumption: there is no hard and fast boundary between the two types of proposition. His claim that 'all observations are theory-laden' is best made clear with an example. A biologist who claims that he can 'observe' a highly convoluted part of an internal cell-membrane, and who claims that this observation refutes the claim of histologists that the 'Golgi body' was a particular and discrete cell organelle, is not making an observation statement in the sense of making a claim based on his unaided senses.

It is only in the light of the theory of electron-optics which determines the construction of the electron-microscope through which the cell is observed that the biologist can claim that he has refuted the histologist's claim that the 'Golgi Body' is a discrete organelle by 'verifying' the 'observation' statement " This part of the cell consists of a highly convoluted part of the endoplasmic reticulum and is connected to the rest of the cell membrane."

Even for 'un-aided' observations the same problem arises, for we can argue that the truth value of the 'observation' statement can be determined only in the light of a physiological theory of the functioning of the human sense organs.

Secondly, this simple form of falsificationism is untenable because it assumes that the truth value of the 'observation' statement can be indubitably decided by experiment. It is, however, an indubitable matter of logic that no factual propositions can ever be proved from experiment. The only method of establishing the truth value of any statement (short of incorporating it as an axiom of a formal system) is by valid inference from other true statements. So 'observation' statements remain for ever unprovable from experience.

Once these assumptions have been exposed, it can be seen that the methodology based on this simple falsificationism does not get off the ground. Observational propositions, being to a greater or lesser extent dependent upon theory, and in any case being unprovable and hence fallible, cannot serve as the indubitable basis for the refutation of theories. Not only are all theories unprovable, but also they are all unfalsifiable.

How, then, can we hope to achieve an adequate account of the growth of conjectural knowledge if our attempts at falsification produce not, as was hoped, grounds for rejecting theories as false, but merely the exposure of an inconsistency between statements ?

Popper's answer is by way of a retreat to a conventionalist position - but not the conservative conventionalism of (among others) Poincare,[4] whose conventional decisions made unfalsifiable, by <u>fiat</u>, theories which had proved their worth by a considerable period of empirical success. Popper's conventionalism decrees that the statements decided by agreement are not universal, but singular. These Popper calls 'basic statements' and they have the form of singular existential statements, and every test of a theory, whether resulting in its corroboration or its 'falsification', must stop at some basic statement which we decide to accept. We can distinguish two separate sets of conventional decision procedures; these correspond to the assumptions made by the simple falsificationist methodology outlined above. Firstly we must decide just which statements are basic statements, and secondly, we must decide which of these we choose to be accepted. The first decision is regulated by the following conditions which basic statements must satisfy. Firstly, the logical conditions which must hold good. (1) From a universal statement without statements of initial conditions, no basic statement can be deduced; (2) For any basic statement there must be a possible universal statement which are incompatible; (3) Basic statements must have the form of singular existential statements. Secondly, they must satisfy the material requirement that the event which the basic statement tells us is occuring, or will occur, at a particular place must be observable - that is to say, if an observer is situated at a suitable place in space and time, he must be able to agree or disagree with the basic statement so tested. Moreover, any observer so placed must reach the same verdict.

The second set of conventional decisions corresponds to the original assumption that statements can be proved from experience. Popper's methodological falsificationism introduces asecond set of conventional decision procedures which decide which of the acceptable basic statements are, in fact, accepted.

Both these sets of decision procedures are fallible, but as they serve in Popper's methodological falsificationism not as the grounds for the proof of falsity of theories, but only as the ground for methodological rejection, this does not matter. A basic statement which we accept for the purpose of one experiment may, in the context of another experiment, be considered problematic. From a logical point of view, there is no reason why we should consider the acceptance of any basic statement as finally, conclusively, decided. Any basic statement can, in its turn, be subjected to tests, using other basic statements which can, with the help of theories and statements of initial conditions, be deduced from it. We stop examing whether a particular basic statement has been correctly accepted not for any logical reasons, but only because that basic statement is of a kind especially easy to test and beaause a large number of observers have reached agreement.

Popper has thus reduced the problem of 'falsifying' theories to the problem of 'falsifying' basic statements. If we define 'potential falsifier of a theory' $=_{df}$ 'a member of a non-empty class of basic statements which are forbidden by that theory', the list of accepted falsifiers is given by the conventional verdict of the empirical scientist.

There are, however, more conventional decisions= associated with Popper's falsificationism. When a scientist conducts an experiment, he must regard all the theories he uses as fallible, and yet he tentatively 'accepts' those which he regards as not being under test. The interpretation of this observational data may involve the application of a (fallible) theory, in the light of which he decides to accept the data as a basis for a decision to accept a potential falsifier of the theory under test. For the purposes of this experiment he relagates this 'interpretative' theory to unproblematic background knowledge, which is tentatively accepted during the test.

This demarcation between background knowledge and problematic theories has a role also in a third type of conventional decision associated with Popper's falsificationism. Many theories do not specifically forbid any observational statements describing any particular observational state of affairs. As they stand their class of potential falsifiers is empty. For example, given a set of initial conditions concerning a particular planet, Newtonian gravitational theory does not, by itself, forbid the planet moving in a helical orbit - or indeed any configuration of orbit you wish - <u>provided</u> that the necessary perturbing forces causing the deviation from the standard ellipse are present. The motion of a planet in a helical orbit is only forbidden by a conjunction of a basic statement describing a particular event (the planet's elliptical motion) and a universal non-existential statement to the effect that nothing occurs to upset the <u>status quo</u>.

This appended clause (the <u>ceteris paribus</u> clause) can be tested by assuming that there <u>are</u> disturbing influences - that 'all other things are <u>not</u> equal', and if many of these assumptions are refuted then the <u>ceteris paribus</u> clause will be regarded as well-corroborated. The decision, however, to <u>accept</u> the <u>ceteris paribus</u> clause will be regarded as a purely conventional one; but once such a decision has been made it can turn what was previously a merely as-yet-unexplained anomaly of the theory into a potential falsifier.

Although this third type of conventional decision can never be justified, it must be taken if theories such as Newton's are to be regarded as falsifiable, and hence, according to Popper's demarcation criterion, scientific.

Yet one more type of conventional decision must be admitted to Popper's falsificationist methodology. If we are to class among falsifiable theories those which are <u>probabilistic</u>, we need a fourth type of decision by which we can interpret statistical evidence as being inconsistent with a probabilistic theory.

These, then, are the conventional decisions which are needed for us to escape the charges levelled against the simple falsificationist programme. In escaping these charges we have radically changed the nature of the concept of falsification, from a logical proof of falsity to a fallible decision to interpret a theory as false. Any methodology so constituted must guard against the criticism that it is perfectly possible to save any theory from this conventional falsification by re-assessing any of the conventional decisions which have to be taken to expose that theory in isolation to the arrow of the <u>modus tollens</u>. For example, if we claim that Newton's theory has been refuted by the accepted potential falsifier "The advance of Mercury's perihelion is 43" per century", this decision to consider Newtonian theory as falsified could be changed by re-assessing our acceptance of (1) the basic statements describing the initial conditions, (2) the observational statements describing the motion of the perihelion, or by (3) re-assessing the <u>ceteris paribus</u> clause - the assumption that there are no other relevant factors in play.

Popper admits the justice of this criticism but, he claims, the empirical method which he characterizes is one which precisely excludes the unsatisfactory ways of avoiding falsification that the critic insists are logically possible. The aim of this empirical method is not to save the lives of untenable systems "but, on the contrary, to select the one which is by comparison the fittest, by exposing them all to the fiercest struggle for survival."[5] We can consider Popper's philosophy of science as composed of two (not necessarily distinct) elements. The first is a conventionalist method for rejecting a theory by falsification. (From now on I will refer to this type of falsification as 'falsification$_P$'). The second is an empirical methodology which ensures that we cannot save an old theory by unsatisfactory means. This empirical methodology reaches its most refined form in Popper's later publications, but many of the ideas were present from the start. To prevent these unsatisfactory methods of saving a theory, "*ad hoc* theories" or "conventionalist stratagems" in Popper's terminology, this empirical methodology stipulates that any new theory should satisfy three requirements:

(1) A new theory should be more highly testable. In other words the class of potential falsifiers of the new theory should contain all the members of the class of potential falsifiers of the old - *and more*.

(2) The new theory should be independently testable. In other words it must predict new, testable, consequences which have, so far, not been observed.

(3) The new theory must pass some new and severe test. In other words some of the independent consequences of the new theory must actually be observed.

These requirements effectively forbid any *ad hoc* saving of an old theory.

This brief outline of Popper's theory of method although by no means a complete exposition, will serve to illustrate the points of attack of Lakatos' criticisms. Lakatos calls a falsificationist programme based on the elimination of $falsified_P$ theories, 'naive methodological falsificationism' (I will abbreviate this to 'NMF'). (Note that this designation only applies to Popper's theory of method with certain reservations. q.v. p106)He points out what he considers to be several unsatisfactory features of NMF, and proposes to replace it with a new, 'sophisticated methodological falsificationism' - which I will abbreviate to 'SMF ! I wish now to introduce just a basic outline of SMF so that the ground may be cleared for a more detailed discussion later on.

Basically, Lakatos' idea is to build into the conditions required for the elimination of theories, just those conditions required by Popper to ensure that new theories are not arrived at by a conventionalist stratagem. This he does by a reformulation of the demarcation principle and the introduction of new requirements for 'falsification'. In NMF the demarcation principle reads:- "Any theory which can be regarded as experimentally $falsifiable_P$ is acceptable, or scientific."
In SMF:-

> "A theory is 'acceptable' or 'scientific' only if it has corroborated excess empirical content over its predecessor (or rival), that is, only if it leads to the discovery of novel facts." 6

In NMF, a theory is falsified by a well-corroborated falsifying hypothesis which is interpreted as conflicting with it. In SMF, a theory (T) is regarded as falsified:-

"If and only if another theory T' has been prop-

> osed with the following characteristics. (1) T' has excess empirical content over T: that is, it predicts novel facts, that is, facts improbable in the light of, or even forbidden, by T; (2) T' explains the previous success of T, that is, all the unrefuted content of T is contained (within limits of observational error) in the content of T'; and (3) some of the excess content of T' is corroborated." 7

One effect of the introduction of these definitions is that in SMF a theory is never falsified$_P$ before its elimination. We never have to commit ourselves to a conventional decision to the effect that the mistaken part of the system under test is the theory itself. In fact, when presented with an inconsistency between the prediction of a theory under test and an observational statement;

> "we do not have to decide which ingredients of the theory we regard as problematic and which ones as unproblematic; we regard all ingredients as problematic in the light of the conflicting accepted basic statements and try to replace all of them. If we suceed in replacing some ingredient in a progressive way (that is, in such a way that the replacement has more corroborated empirical content than the original) we call it falsified$_L$." 8

In Popper's falsificationism the ground for declaring a theory false$_P$ is an interpreted conflict between theory and an accepted observation statement. * To __accept__ a theory, however, it must be assessed in relation to the theory it is replacing. In other words, to appraise the intro-

* This statement will need some elaboration. See p.106

duction of a new theory we must consider, not theories in isolation, but sets of theories. In SMF both accepting <u>and rejecting</u> theories calls for appraisal of sets of theories. Indeed the demarcation principle of SMF makes the notion of a single 'isolated' theory unscientific.

What are the benefits of SMF over NMF ? Firstly we can dispense completely with the third type of conventional decision needed for the NMF methodology. The <u>ceteris paribus</u> clause is just taken as one of the problematic elements of the theory we are considering. As we now are not in the position of having to falsify$_P$ the theory under test, the decision to the effect that there are no disturbing influences, need not be taken. If the decision to assume the falsity of the <u>ceteris paribus</u> clause leads to new and corroborated predictions, then, and only then, can we say that the old clause was false$_L$. If, on the contrary, we find that the decision to replace the <u>theory</u> with a new one leads to new, and corroborated, predictions, then, and only then, we have falsified$_L$ the old theory.

Secondly, Lakatos claims that we can reduce the reliance which we place in the first two types of methodological decision. This is achieved by the introduction of an <u>appeal procedure</u>. If, for example, we are using a particular observational theory, in the light of which we decide to give the truth value 'true' to a potential falsifier$_p$ of a theory under test, the sophisticated methodological falsificationist would not deny the possibility of a change in this assignation. The theory under test would at one moment be f falsified$_p$, and at the next moment it would have been reprieved. This possibility of reprieve of the theory by change in the truth value of potential falsifiers is impossible under a NMF programme - when a theory has been falsified$_p$, it must be ruthlessly rejected. And yet, Lakatos claims, the actual history of science provides many examples where this type of reprieve has taken place. SMF describes the scientific method as actually practised by scientists more faithfully than NMF.

Because SMF reduces the number of, and the reliance which we place on, conventional decisions, each of which we recognise to be fallible, Lakatos claims that it is more satisfactory. He argues that it can only be to the good that we have reduced the number of occasions when we are liable to be in error. As, it is implied, the new

SMF methodology has all the positive virtues of the old NMF, in giving a satisfactory account of the growth of scientific knowledge, we should accept it as a more rational theory of criticism.

This completes my introduction of the methodology of Lakatos. I shall be returning to it, and to an examination of the questions which it raises, in the main body of the essay.

CHAPTER ONE

SECTION A

This essay will be concerned with the use of the terms 'competition', 'preference' and 'betterness', together with their numerous cognates, which occur in the vocabulary which scientists and philosophers of science use in appraising scientific theories. Each of these terms has many different uses. We can often get some more or less clear idea of the nature of a particular use of a term by subjecting a statement in which the term occurs to some sort of analysis. We can frequently differentiate between different uses of a term by examining the grounds which we have for using a term on a particular occasion. If we find that a different analysis is required to determine the grounds for use of one occurence of a term from the determination of the grounds for the use of a second occurence, then we may reasonably suspect that the term has different uses in its two occurences. (This is not to say that we might not have the same grounds for two distinct uses.)

It seems that, in the present inquiry, there are two distinct possibilities. It may be that the determination of the grounds for a use is analysable in a purely logical manner, or it may be that our analysis specifically requires extra-logical considerations. In other words, we may either be able to determine the grounds for a particular use solely by a consideration of certain logical relations which obtain between the statement under analysis and the statements which have these relations to it, or we may find that the grounds for a particular use depend upon some other factors apart from logical relations between statements. These extra-logical factors are usually expressed in the form of regulative methodological principles.

We do not, however, need to examine the grounds for the uses of 'choice' <u>and</u> 'preference' <u>and</u> 'betterness'. These terms are so related that once we have established the grounds for the various uses of 'betterness' a simple schema will (with one or two quite unexceptional assumptions) provide the grounds for the uses of the other terms.

These terms are related in the following manner. All forms of preference stand immediately in a relation to betterness. There are two forms of this relation: (1) we may prefer a certain thing or state of affairs to another <u>because</u> the former is, or is thought to be, <u>better</u> ina a certain respect than the latter. Here we can offer as a ground for a particular preference the conjecture that one state of affairs is better, in a certain respect, than another; (2) we may prefer one thing or state of affairs to another simply because the first is <u>liked</u> more than the second. Here we can not offer as a ground for our preference our liking it more. The first form of preference is <u>objective</u> and the second, <u>subjective.</u> Any attempt to twist the second form by (for example) claiming that we like one thing more than another 'because it is better in soliciting our approval' just results in our saying 'I like it because I like it.'

Thus we can in one case offer the grounds for betterness as indirect grounds for preference, and in the second case where we have no grounds for betterness we can (obviously) offer no such grounds for preference. It is because of this problem that we introduce the first of our assumptions. If we wish to offer, as the grounds for the uses of 'preference' in the scientists' appraisals, a conjecture of betterness, then we must eliminate from those appraisals any subjective use of 'preference'. We shall assume, therefore, that any appraisal based on such a subjective preference is not a <u>legitimate</u> appraisal - it is a sort of psuedo-appraisal, because no grounds can be given for the use of the term. A subjective preference can be revoked and reversed without the charge of inconsistency being applicable, at least not <u>logical</u> inconsistency.

For similar reasons there are some uses of 'betterness' that we must outlaw from critical appraisals. These are uses which do not seem to be in any way related to preference. One such use is exemplified in the statement: "Bloggs is a better chess-player than Jones". Betterness in performing some given task is not connected with the notion of preference unless we add a particular goal or end. If we do add such a goal then the connection with preference becomes clear; if <u>we wished to select a good chess team</u> (the goal) then we would prefer to have Bloggs as a member of the team.

Finally the relation between 'choice' and 'preference'. Normally, faced with two things or states of affairs, we choose the one that more nearly satisfies our goal-directed preference. If a person does not so choose, we suspect the presence of some external compulsion, or we doubt his sanity.

This progression, by which the grounds for betterness of a use of 'betterness' generates subsequently the basis for a corresponding use of 'preference' and 'choice' characterises the method by which we can justify the preferential choice of one theory over another in a competition situation.

So, far, my use of the terms 'theory' and 'competition' have been rather vague. It would be useful to clarify these before we go much further. First of all, 'a theory'. It would be an herculean task to give a set of criteria by which we could separate from the body of scientific knowledge just those parts which, at one time or another, some scientist has referred to as 'a theory'. Let us be content to state some conditions which are at least necessary. A theory consists of a set of statements at least one of which is a universal statement, in which no proper names occur. From this universal statement it must be possible to deduce singular non-existential statements which are observation statements i.e. for which there exist some technique such that anyone who has learned this technique can decide whether or not this observation statement is acceptable. Now, as nothing observable <u>at all</u> follows logically from simply a universal statement, the theory must also contain another set of singular statements, those describing initial conditions.

This schema would be satisfied by the following 'theory'. Universal statement: "All swans are white". Initial condition: "This is a swan". Deduced non-existential statement: "There is not, in this particular place, a black swan". Technique: Look and see. Clearly we need to add some further conditions if we are to restrict that which we characterise as a theory to the sort of thing we might find in the pages of a physics text-book. We have, however, given all the fundamental logical requirements, so far as the structure of the component statements is concerned.

We *can* stipulate further conditions, but the difficulty is to make them at all precise. If we, as is customary, claim that a theory *explains* (i.e. can be used to give a causal explanation of) or *accounts for* the events described by its deductive predictions, then it is usually the case that 'text-book' theories can be used to account for much more unfamiliar facts. Yet we would not want to make this a necessary condition for all theories.

Again, it is claimed that a theory should have been arrived at in an attempt to solve some 'deep' and 'interesting' problem. And yet we would not think a theory which solved the unified field problem any less 'a theory' if it had been proposed as a joke. Although a situation in which a scientist is grappling with a difficult problem makes a good spawning ground for theories, we cannot demand it as a necessary condition for a set of statements constituting a theory.

For the classical empiricist the only <u>genuine</u> theory was one which satisfied the conditions of axiomatization, and, moreover, where the interpretation of the axiom system was intuitively certain, of self-evident. As Kantians discovered, on the successful introduction of empirical non-Euclidian geometries, this is asking too much.

Let us rather, then, accept theories of the 'All swans are white' category as 'philosophers' theories', but label them, for reasons which will be given later, bad scientific theories.

I wish now to turn to the distinction between observational and explanatory theories. An explanatory theory is one which we can use to deduce observational statements. An observational theory is one, in the light of which, we regard observational statments as acceptable or not. Thus <u>all</u> theories are explanatory theories in the sense that observational statements can be deduced from them. A theory is classed as 'observational' only relative to a particular experimental context. For example, in the context of a experiment in radio-astronomy, where we are testing (say) some aspect of Newtonian gravitational theory by 'observing' the motion of a radio-star system, we interpret the experimental data in the light of the (observational) theory of the propagation of electromagnetic radiation. In another experiment the theory of the propagation of electro-magnetic radiation might become the explanatory theory and we interpret the experimental data (say, observation of interference patterns) in the light of some theory which in this new context now becomes the observational theory. (Here, for example, we could rely on the theory of heat conduction, in the light of which we could interpret data concerning heat sensors as indicating points of maximum energy in the interference pattern).

I propose to introduce the term'theory system' to indicate the conjunction of an explanatory theory and those observational theories which are necessary for the assessment of the acceptance of the basic observation statements involved in any test of the explanatory theory. These observation statements include statements of initial conditions and observational statements of prediction. It may be argued that this makes just what constitutes a particular theory system dependent upon the person who is making the acceptance decisions. One scientist may interpret the evidence in the light of onw observational theory and another scientist may use a second. Alternatively, one may demand that, to be justified in accepting a basic statements in the light of an observational theory, he must first have reason to believe that this observational theory itself has been well-corroborated, and accordingly would demand that its deductive predictions have been tested. In doing this he would have to have recourse to another observational theory or theories in the light of which he can interpret the evidence for the first. Are we to class these other observational theories in our original theory system ? This would lead to an infinite regress at worst, and at best it would necessitate the inclusion of a vast number of theories in a viscious circle, each theory depending on another for the acceptance of its observation statements. This problem is associated with what Popper has called the 'relativity of basic statements': his answer being that we continue the chain of acceptance-dependency until we arrive at basic statements which are (1) very easy to test, and (2) generally accepted. In other words, there must be a very large degree of inter-subjective agreement as to their acceptance. This provides a methodological principle by which we can limit the regress. We class in the theory system all those observational theories up to, but not including, those whose predictions are observation statements which satisfy the above conditions.

There is one more element of the theory system, a falsifying hypothesis. Faced with a <u>single</u> instance where our interpretation of an experimental result demands that we do not accept an observation statement which is a non-existential singular statement deduced from the explanatory theory, we do not consider this alone to be sufficient to count against the theory system. We frame a hypothesis that this instance is typical of a general regularity and was not due to 'accidental' error incurred during the experiment. Only when this 'falsifying hypothesis' has been well tested do we agree that our expectations have been disappointed. The observational theories which we need to test this falsifying hypothesis may or may not be the same as the ones in the light of which we 'provisionally' rejected the observation statement. If they are not the same, we must include these additional observational theories in the theory system.

A theory system is the smallest expirical epistemological unit of which we can say on the basis of our experience 'It is inconsistent'. This notion of the inconsistency of a theoretical system is closely associated with falsification$_L$. Because we are not logically justified in claiming that the inconsistency is due to one <u>particular</u> component of the system (one which we can identify), our critical appraisal is directed at the system in toto. When we $falsify_P$ an explanatory theory and claim that the universal generalisation(s) is/are in error, we rely on certain logically unjustifiable assumptions by which we regard the explanatory theory as being responsible for the inconsistency.

Expressed in another way, the two approaches of naive and sophisticated methodological falsificationism could be differentiated by the distinct concepts of the empirical basis of science which they employ. Both approaches associate the empirical nature of science with the growth of scientific knowledge, but the naive methodology contains an additional element. This is the idea that scientific theories prove their worth in competition with the 'facts' (or at least what we choose to interpret as the 'facts'). The empirical nature of an explanatory theory shines without blemish once it has been $falsified_P$. On the other hand, in SMF we have no concept of competition with the 'facts'. When one theory system has successfully replaced another (i.e. when the first has been $falsified_L$) we, apparently, cannot argue from this to reach any conclusion that the explanatory part of the theory system fell short of its task of truly predicting the 'facts'. In SMF the concept of the empirical basis of science is associated only with the idea of its growth.

I shall deal later with the possible justification we have for employing methodological decisions which enable us to falsify$_P$ explanatory theories and the associated idea of competition of an explanatory theory with the 'facts'.

Before we leave this characterisation of theories, I must add one word on the matter of consistency. We require of a theory that it is consistent in the sense that it shall not be possible to deduce contradictory observation statements from it. We require, more specifically, that it shall not be possible to deduce from the universal statements and the same set of initial conditions, two observation statements which are contradictory.

Let us turn to clarification of my intended uses of the term 'competition'. I shall use the term 'competition' in the context of the assessment of theories, to describe a situation in which there is a possibility of preferential choice between two theories. If this preferential choice is to serve as the grounds for the elimination of the non-chosen theory, then we require that the two theories be comparable, and I shall give the conditions required for comparability in the next section. In general, we shall only be concerned with choice between comparable theories. For now let us accept that when the results of a competition are obtained we will be able to give a reason for the choice of one theory over another.

A competition involves the notion of a primary goal (the aim of the competition), and a set of rules which regulate fair play. In the present context, I shall call these rules the 'rules of comparison'. By following the rules of comparison we will be able to say with justification that one theory is better than the other, if indeed it is. The primary goal of the competition is to endevour to rate as highly as possible for each competing theory, the notion of betterness which we are assessing. In choosing between competitors, we will need a <u>secondary goal</u>. We can make this concept of a secondary goal clear by a simple analogy. Consider a typical, every-day competition - say, a football match. The primary goal of a football team is to score goals (using 'goals' in a rather different sense). At the end of a particular match we can say that one team is better than the other <u>because</u> it scored more goals,and scoring goals was the aim of the competition. However, this does not, as it stands, give a reason for the preferential choice of the winning team. However, the England Team

Manager, (sitting on the touch line), who has the job of selecting the strongest team for the European Cup, has, in the light of this <u>secondary goal</u> of choosing the strongest team, a reason for the preferential choice of the winning team.

Now, no analogies are perfect (and we shall see that this one quickly breaks down), but there are some obvious points of similarity. We choose the theory which has 'won' the competition because that choice takes us nearer to achieving our secondary goal. A typical formulation of a secondary goal would be 'to achieve a more satisfactory explanation of whatever strikes us as being in need of an explanation'.

Many separate secondary goals may be relavent in the choice made as a result of the outcome of a <u>single</u> competition. If, for example, the outcome of a competition was that theory T was better than theory T' in respect of having a higher degree of falsification, then we would preferentially choose theory T in the light of the following secondary goals: (1) To arrive at theories which are more highly falsifiable, (2) To arrive at theories with a higher empirical content, (3) To arrive at theories which are (in a Popperian sense of the word) simpler.

Many other secondary goals will not be <u>relevant</u> to this particular competition. We need not, for example, in the above case, necessarily choose T over T' if our only secondary goal is to arrive at theories which have been more highly confirmed by the available evidence. Indeed in this case we have no grounds for making any choice at all.

I will call secondary goals <u>compatible</u> if in each competition to which they are relevant the same preferential choice is made in the light of each of those goals individually. I will call two secondary goals <u>incompatible</u> if there exists a particular competition such that in the light of one goal we choose one theory and in the light of the other goal we choose the second theory.

In general, the intuitive notion of the 'progress of science' ensures that the secondary goals which a scientist strives for are compatible, but there are exceptions to this. Some secondary goals have a self-evident justification. We do not need to argue in favour of the adoption of such goals as 'To arrive at theories of an ever-increasing explanatory power'. Reasons given for its adoption might be 'Well, that's what science is about' or 'What would it be like to have as a goal '<u>Nòt</u> to arrive at theories of an ever-increasing explanatory power''.

Other secondary goals are more in need of justification and require some explicit assumptions before they can be regarded as tenable. Some philosophers of science have proposed the secondary goal 'To arrive at simpler theories', where '<u>simpler</u>' is used in the (rather obscure) sense of 'theories describing a simpler state of nature'. But until they can provide some sort of satisfactory explanation which makes explicit why the acquisition of theories chosen in the light of such a goal would be advantageous, it seems dubious to claim that we should try to achieve it.

Essentially, I believe, the adoption of all secondary goals depends upon some pragmatic consideration. The adoption of a secondary goal implies that we believe that the achievement of this goal will be of some pragmatic value, either in the sense that the results of the selection procedure will be such as to provide us with theories which are more useful in adding to our knowledge or more useful in adding to our technology.

The various grounds which we can give for claiming that one theory is better than another are intrinsically related to the secondary goals which we deem relevant to theory choice. It is no accident that we choose the theory which we claim, in the light of the rules of comparison, is better. In general, the particular rules of comparison which we apply are just those rules which assign betterness to the theory which can be used to approach the secondary goal more nearly. It would seem then that if a notion of betterness which we employ is to be considered a satisfactory notion, we must always be able to provide the explanation as to why better theories are more likely to enable us to achieve the secondary goal. Sometimes, however, this explanation would be gratuitous, if the secondary goal were sufficiently 'self-evident'.

Secondary goals have varying degrees of universality but any attempt to provide a method of ordering, has to rely on intuition. At one end of the scale we have goals which are <u>particular</u>, such as 'To predict more accurately the moon's relative motion' and at the other end of the scale we have <u>general</u> goals, such as 'To arrive at the most probable theories'. From now on, I shall be considering only those goals which are general.

In the light of a general goal, we can make a choice between <u>any</u> two competing theories. We can thus interpret these secondary goals as general principles of scientific methodology. Such principles are, however, <u>useless</u> unless we can lay down a set of constraining conditions under which we require them to operate. There is no point, for example, in choosing the simpler theory (in some well defined sense of 'simple') if the chosen theory does not explain our observational data at least as well as the more complex theory. Although we may strive to reach the goal 'To arrive at simpler theories' we need to operate under constraints which prevent us from choosing theories which do not have also certain desired characteristics. Frequently these constraints are expressible in the demand that several different notions of betterness are satisfied simultaneously.

Now let us return to the analogy of the football match and point out its inadequacies. In the match itself, the only rules of comparison which determine which team is the better, are those by which we decide which team has scored the most goals. The situation in theory competition is more complex. We have several different <u>methods</u> of comparison (q.v.) which we can employ. The analogy would be closer if we <u>also</u> chose between the teams in the light of the football teams' style, or the respective heights of the players in each team, etc., and for each choice we may need a separate method of comparison. In theory competitions, I believe that there are at least three different methods of comparison which are essential to any philosophically adequate account of scientific progress. These I will list and examine separately. (1) _A priori_ comparisons, (2) Comparisons of a theory with facts, and (3) Comparisons based on the notion of progressiveness.

(1) _A Priori Comparisons_.

By '_a priori_' I mean 'prior to any empirical testing of the competing theories'. In this method we assign to the theory a value of some measure function, or set of statements which have a certain logical relation to the theory. These must be determinable without empirical testing. We can then, compare theories by comparison of these sets of statements or of the respective function values.

Consider as an example of this type of competition, the case where we assign to theories a measure function, the value of which depends upon the number of parameters in some particular statement of each theory. If we have a set of experimental data which can be represented as points on a graph, then there are infinitely many curves which can be drawn through these points. In many cases the equation which represents these curves can be considered as statements of a theory. Let us consider the simple case where each of the many possible equations is a statement of a separate theory, and also where each of these equations generates a curve which fits the experimental points to some desired degree of accuracy. We can compare these theories by inspection of the number of parameters in these equations. The rules of comparison would be, '(1) Add one to the score for each theory for every occurence of a parameter in the equation' and '(2) The better theory is the one with the lower score'. The grounds for the use of 'better' in this case can be analysed purely in terms of a formal analysis of the equations. In other cases of a priori comparison, we find that a requisite analysis cannot be supplied in purely formal terms. Consider, for example, Popper's notion of degree of falsifiability and how we can analyse the grounds for the use of 'better' in comparisons based on this notion.

If we have a class of statements, such that a certain conjunction of some number, (d + 1) of these statements, ~~such~~falsify a theory t, but any conjunction of d statements of that class can not falsify the theory, then we call d the dimension of that theory with respect to this class of statements. (This class of statements is called the field of application.)

We can now put forward the following rules of comparison: (1) Assign to each competing theory t_1, t_2, its dimensional number d. (2) A theory t_1 is better (i.e. has a higher degree of falsification) than the theory t_2 if the dimensional number of t_1 is less than the dimensional number of t_2.

The secondary goal (to seek the most highly falsifiable theories) which directs our choice of t_1 must operate under the constraining condition that both dimensional numbers are measured with respect to the same field of application. As in the previous example a comparison is achieved which is not dependent upon any empirical testing of the competing theories. However, in the second example the analysis of the grounds for betterness in respect of degree of falsifiability cannot be conducted in purely formal terms. We rely on the notion of falsification$_P$ which, as has been explained earlier, depends upon several conventional methodological decisions. It is, however, fair to say as Popper does, that the analysis is 'largely formal'.

These are two examples which I hope will make clear what I mean by 'a priori' comparison. (I will call these typr 1 conparisons.)

(2) <u>Comparison of Theory with Facts</u>.

In the above cases, the comparative procedure was to assess each theory separately, then to compare competing theories by an examination of the assessments. These assessments were independent of any evidence for or against the theory. In comparison of theories with facts, we still assess the theories concerned separately, but that assessment is dependent upon the outcome of certain experiments.

Of course, I am using the word 'fact' with a certain degree of licence. To say just <u>exactly</u> what a fact <u>is</u>, is rather problematic. However, I can clarify the notion of comparison with facts by specifying the type of secondary goal which is relevant. These are goals which involve the notions of arriving at theories which we believe are true or false, or probable or improbable. In other words, the use of these goals involves the idea that we should <u>justify</u> our acceptance or rejection of theories and that this justification involves giving reasons why we believe them to be true or false, or to have a certain degree of probability. There are, of course, enormous difficulties present in this way of speaking, and it may even be reasonably claimed that we can give a satisfactory account of what constitutes progress in science without reference to truth or probability. Nevertheless I think that unless this account can be as interpreted to give reasons why we believe that our accepted theories are at least nearer to the truth, they cannot be considered as giving a philosophically adequate account of scientific progress. If this is so, then our assessment of theories must include at least one comparison of theories with the facts. (I will call these comparisons type 2 comparisons.)

A description of type 2 comparisons which have been suggested to determine which theories can be used to achieve these secondary goals would include, once again, formal and informal methods of analysis.

Consider Carnap's original programme (prior to 'Testability and Meaning') of the probabilistic confirmation of theories, by which he claimed theories could be compared by the degree to which they had been confirmed by the facts. If this programme could be achieved we would have a means of assigning to the competing theories the value of a measure function which Carnap called 'the degree of evidential support'. Just what this value is depends upon experimental evidence that has not yet refuted the theory, but has been in accord with its deductive predictions.

We can compare two theories by inspection of the values of this measure function. The rule is: 'The theory which has the highest value of the measure function is the better', and we choose the better theory in the light of the secondary goal, 'To arrive at theories which are more highly confirmed by the evidence' which Carnap interpreted as equivalent to 'more probable theories'. According to Carnap, the grounds for this use of 'better' can be analysed formally, in terms of an 'analytic' (i.e. infallible) inductive logic. Given <u>any</u> theory and the relevant evidence for that theory, we follow a formal procedure to arrive at the particular value of the measure function. Carnap's secondary goal is under several constraints, one of which is that the measure function must satisfy the requirements of an intuitive concept of 'confirmation'.

As an example of an informal type 2 comparison, we can cite comparisons based on the notion of falsification$_P$. Here we can conceive the immediate goal of the competition as an attempt to elicit the verdict 'the theory is false$_P$', in the face of certain experimental evidence. Two different uses of 'better' are associated with this comparison. We say that (1) if for two comparable theories, one elicits the verdict 'it is false' but the other does not, then the theory which elicits that verdicts is $better_1$ than the other, and (2) if for two comparable theories, one elicits the verdict 'it is false' but the other does not, then the theory which does <u>not</u> elicit that verdict is $better_2$ than the other. By pointing out these two uses, we can clarify the confusion which some commentators have had in interpreting Popper. Why should we think that theories which are false are better, when we obviously are engaged in a search for truth? The fact is that two <u>incompatible</u> secondary goals are relevant to this competition. (1) To arrive at theories which we decide to call false, and (2) To arrive at theories which, we conjecture, are nearer the truth. In the light of goal(1) we choose that $false_P$ theory, and goal (2) we choose the one <u>not</u> thought to be false. Later, I will discuss in more detail the need for these separate goals, and why it would be beneficial to achieve them. Briefly, I believe that it is only by achieving goal (1) that we have the possibility of achieving goal (2).

It is obvious that there is no formal procedure for analysing the ground for the use of 'better' in this comparison. It is a matter of empirical experiment whether the competing theories pass or fail the test.

(3) <u>Comparisons Based on the Notion of Progressiveness</u>.

The third method is in some respects, an amalgam of the other two. One <u>a priori</u> comparison we mentioned was where for two theories we assigned classes of statements which stood in a certain logical relation to two theories respectively. Then we could compare these theories by an examination of these two classes. An example of this would be the comparison of empirical content (in a Popparian sense) of two theories by seeing if the class of potential falsifiers of one theory was a sub-class of the class of potential falsifiers of the second theory. If our secondary goal involves some notion of <u>excess</u> betterness (say, empirical content) and we interpret this as being measured by the statements we assign to the theories, then obviously, the comparison must be made with reference to <u>both</u> sets of statements.

In the second method of comparison, we can also think of statements being assigned to the theory, but in this case we can only decide whether the relation between theory and statement holds, by conducting an experiment, (e.g. the relation between statement and theory 'is an empirically verified deductive consequence of'). Here we compare theories with respect to their performances in accounting for the _same_ statements. In other words, we examine the relations which hold (or do not hold) between two theories and the _same_ statements.

In the method of comparison which I call 'comparisons based on the notion of progressiveness' (type 3 comparisons), the rules of comparison stipulate the conditions (1) the classes of statements assigned to each theory are not coextensive, (2) whether a particular logical relation holds between the statements which make up the complement of these classes and one of these theories, depends on the outcome of experiments.

An example of type 3 comparison is the falsificationist comparison of excess corroboration. Here a theory is said to be better than another if it successfully predicts novel facts. As novel here means facts not predicted by the other theory we need (1) to establish that the classes of deductive consequences of the theories are not coextensive (which we can do by an _a priori_ comparison) and (2) to establish by experiment that the relation 'is a successful prediction of' holds between the statements which are a consequence of one theory only and that theory. This type of comparison cannot be used to assess theories individually, but only with respect to another theory. _A priori_ and type 2 comparisons can be used _either_ to assign some 'absolute' assessment to an individual theory _or_ to assess some difference in degree of betterness between two theories.

These are the methods of comparison appropriate to assessment of theories and theory systems. I believe that any legitimate assessment falls into one of these methods. What I hope to do, is to demonstrate that, if we are to provide an adequate account of scientific progress we need to make comparisons by each of these methods. This can be done by showing that the account of scientific progress is inadequate if it lacks instances of any one of these methods, and accordingly we can show that we must include at least one representitive of each of these methods.

The particular members of each class we choose are, to some extent a matter of personal preference. However, _this_ preference also is goal directed, and I shall show that some members which have, historically, been selected cannot be used to achieve the implied secondary goals. Thus we must restrict our choice of comparisons to those which have not been demonstrated to be unsatisfactory.

SECTION B.

In this section I wish to consider the conditions which I will require for the possibility of comparison and to introduce the terms '<u>comparable</u>' and '<u>eliminate</u>' both of which I will use in a technical sense. (Here I shall give only a necessary condition for elimination).

So far our analysis allows for such choices as, 'Choose Dalton's Atomic Theory over Quantum Mechanics because it is simpler'. I would not want to revent anyone from making such a choice - we could approach many secondary goals by doing so, e.g. 'To arrive at theories more suitable to teach to a class of twelve year old children'. For such a choice to be possible the theories must be commensurable with respect to the particular notion of betterness which is being assessed. This is a general requirement. If we are to discard one theory then the theories so assessed must be incompatible with respect to the secondary goal which we hold. Tis also is a general requirement. However, such a secondary goal as 'To arrive at theories more suitable to teach to a class of twelve year old children' is not really of interest to a scientist <u>qua</u> scientist nor to a philosopher of the growth of knowledge. To restrict the comparisons to 'interesting' ones we shall require the chosen theory to be <u>comparable</u> with the one not chosen.

The concept of 'comparability' (in this technical sense) depends upon the notion of an <u>explanatory</u> <u>field</u>. The explanatory field of a theory is the class of accepted basic statements which can be deduced from that theory with the aid of a certain body of background knowledge. If is the expanatory field of theory T, and the explanatory field of theory T', relative to the same body of background knowledge, then T is comparable to T' if and only if, either is a proper subclass of , or has the same extension as . Now we can say that if T' is comparable to T, then any other competition outcome which results in our choosing T'is a ground for the <u>elimination</u> of T.

The justification of the introduction of these concepts of 'comparability' and 'elimination', is that we must have some means of satisfying the intuitive requirement that we only discard a theory from the body of scientific knowledge - strike it out of the text-books of the day, as it were - if we have some alternative theory at our disposal which is capable of being used to perform at least as well, if not better, the function of the old theory. The introduction of 'comparability' provides a more precise formulation of 'performs the same function as'; it interprets the function of a theory as essentially a means of providing explanations by deducing observation statements. As a result of a competition we may discard a theory if the rules of comparison of that competition indicate that we assess the other theory as better. Unless, however, the chosen theory is comparable to the discarded theory, we have no grounds for striking the discarded theory from the body of scientific knowledge, i.e. for eliminating it.

Note that the explanatory field of a theory is not co-extensive with the class of observation statements which are permitted by the theory. Elementary particle theory, for example, permits observation statements such as 'At x,y,z,t, (some spatio-temporal co-ordinates) a particle which has the mass of six electrons can be found'. This is a <u>possible</u> state of affairs which is deducible from elementary particle theory, but in fact, so far no such particle has been observed. If such a particle ever were to be found, then the observation statement describing this state of affairs would form part of the explanatory field of the theory, but not until then.

These requirements are tantamount to the requirement that before we con eliminate a theory, we require a type 1 comparison between the theories by means of which we can assess the respective theories as to their explanatory fields. Put in another way, it amounts to a claim that any 'interesting' comparison must be one which is carried out in the light of a secondary goal which operates under the constraint that we are always seeking theories which explain more.

One of the requirements for the possibility of comparison which I have stated is that the theories be <u>commensurable</u> with respect to the particular notion of betterness assessed. We can demonstrate that this is essential by examining the problems raised by Kuhn's dictum: 'Theories framed in different paradigms are (often) both incompatible and incommensurable.'[19] For our purposes, all that we need to understand by 'in different paradigms' is that a fundamental change of somescientific concept has taken place in the transition from one theory to the other. Let us consider the Newtonian and Einsteinian gravitational theories, (which Kuhn claims are both incommensurable and incompatible) and the different concepts of mass which are involved in these two theories. In the Newtonian system, mass is a concept which results from an interpretation of certain universal statements of Newtonian theory. If we were asked to measure what the magnitude of the mass of a particular bady was, we would have recourse to some procedure which is interpreted as involving the measurement of a _force_ which exists between this body and another. If, on the other hand, we were asked the same question in a relativistic context, our measurement of (from a relativistic point of view) mass would require a procedure which we interpret as needing information of the relative velocity of the body to the observer. On a particular (hypothetical) occasion, we might arrive at different numerical values for the $\text{mass}_{\text{classical}}$ (mass_c) and the $\text{mass}_{\text{relativistic}}$ (mass_r) of a particular body. But, Kuhn claims, it is nonsense to say the mass_c is larger or smaller than the mass_r. One can, of course, compare the _numbers_ by which the magnitudes are expressed, but not the magnitudes themselves because of the different concepts of mass employed. The two concepts are incommensurable. But Kuhn goes on to say that the different _theories_ in which these concepts occur are also incommensurable, and it seems _this_ incommensurability is a result of the different concepts involved in each theory which provides us with a different 'gestalt'. Now it is obviously true that the statement "This body has a $mass_c$ of n mass units" is not incompatible with the statement " This body has a $mass_r$ of n + 1 mass units" (where we refer to the same body). But Kuhn wants to go further than this. He appears to argue that the occurene of these incommensurable concepts in theories makes the notion of incompatibility inappropriate to theories which contain these incompatible statements. _If_ this is what he means by 'incommensurable theories' then we can find reason to disagree. It seems very strange to say of Newton's and Einstein's theories that they are not incompatible, i.e. that they _could_ exist side by side in our body of scientific theories, with no means of putting them into competition. We can deduce from Newtonian theory, given certain initial conditions, an observation statement (A): "There is no apparent displacement of the star's position". This describes our observation of a star which is nearly in conjunction with the sun. We can deduce from Einstein's theory the statement (B): "There is an apparent displacement of the star's position". If we assume that the referents and contexts are identical, then statement A is certainly inconsistent with statement (B). However, we may conduct a particular experiment which leads us to <u>accept</u> statement (B), and it would appear to be the case that it would be inconsistent to accept statement (A) at the same time. Is it not sufficient to say that, if two inconsistent statements can be deduced respectively from two theories and we accept one statement and reject the other, then the two theories must, in some sense, be incompatible?

Kuhn's case sould, however, be argued further. Although statement (A) <u>looks</u> as though it is inconsistent with statement (B), in fact, it may be claimed, that this inconsistency is only apparent. The two statements belong to two different <u>domains of discourse</u> and we should not be confused by their grammatical similarity. They are statements made in the light of two different world views and each is relevant only to observations and predictions made from within that world view. Statement (A) and (B) are not <u>inconsistent</u> - they should not be confused as being similar to (C) The man's height is six feet, and (D) The man's height is not six feet, where (C) and (D) <u>are</u> uttered in the same domain of discourse. An attempt at making the relative domains clear, would produce some such ammendment as: (E) "As the $mass_c$ of light is zero, there are no gravitational forces affecting it as it passes through Euclidean space near to the gravitational field of the sun, so the light will continue in a straight line" and (F) "Although the rest $mass_r$ of a photon is zero, its relativistic $mass_r$ is finite, so that as it passess through non-Euclidean space in the vicinity of the sun, whose mass exerts a distorting effect on the curvature of the space around it, there will be a deflection away from what would be the shortest time-path had the sun not been there".

If these domains of discourse are differentiated, then, we can argue, the notion of inconsistency between statements (A) and (B) breaks down - and hence the <u>theories</u> (A) and (B) can once again be considered incommensurable. We seem to have demonstrated that the theories could exist side by side - i.e. that they are not incompatible. But does this argument hold water? It seems to me that although the theories <u>are</u> incommensurable in respect of <u>certain</u> notions of betterness they <u>are</u> commensurable, even on Kuhn's account, in respect of certain others. If, for example, the particular notion of betterness was 'predicting more accurately the mass of a certain body', then we <u>cannot</u> compare the two theories, because the concepts of mass involved are different. Moreover, if the notion of betterness used was 'predicting more observations than . . .' (i.e. a type 1 comparison), then the rules of this comparison could be followed with no reference to the disparate concepts of the two theories. We can, in this case, consider these observation statements to be expressed in a neutral observation language - one which is independent of the language of the two theories - and carry out our comparison in this language.

Hence we can sdve the apparent contradiction in Kuhn's dictum. Certain statements of two theories may be incommensurable, and comparison involving incommensurable statements is impossible. This is, I think what Kuhn means when he claims that theories are incommensurable. But it does not follow from this, that the theories are incommensurable with respect to all notions of betterness, as we have shown. Moreover, although the particular notion of betterness which we considered was 'predicting more observations than . . .', it is obvious that we could also assess betterness in respect of 'the explanatory field of . . . is a subclass of . . .' and so demonstrate the comparability of Einstein's theory with Newton's.

Although I believe that this idea of conceiving different theories to be framed in two different gestalts, is inherently vague and not particularly useful, we can say that _if_ we accept such a view then the _only_ method of comparison which is possible is a type 1 comparison. We can, of course, refute Newton's theory by observation statements made from _within_ one gestalt, and _not_ be able to refute Einstein's theory from within the other. But in this case the possibility (Kuhn would claim) of performing a type 2 comparison is _empty_; to say that we prefer T' to T _because_ T is refuted and T' is not, is to make the illegitimate comparison between 'refuted' in one world view and 'not refuted' in another.

If we accept that such comparisons _are_ meaningful, then we _can_ say that the grounds for our preference are that one theory is refuted but the other not. But we should recognise clearly the limitations of such a comparison which _limits_ the notion of betterness exclusively to 'better in the respect that it has not been refuted'. We are limited to a comparison of the logical relations which hold between the two theories and statements which can be expressed in a neutral observation language without reference to the incommensurable concepts of the two theories.

CHAPTER TWO

In this chapter, I will examine formal methods of comparison, in section A, comparisons based on the notion of inductive simplicity, and in section B, comparisons based on a group of notions of inductive support..I shall adduce reasons why these, and why many formal methods of comparison, must be considered unsatisfactory.

SECTION A

> "Is there a concept of simplicity which is of importance for the logician? Is it possible to distinguish theories which are not equivalent according to their degrees of simplicity?"[10]

> "When the evidence leaves us with a choice among hypotheses of unequal strength, how is the choice to be made . . . simplicity must be taken into account."[11]

It is clear from these references that some concept of simplicity is held to be of great importance in deciding between competing theories, that the secondary goal 'to arrive at simpler (in some sense) theories' is one which we should strive to achieve. Before we examine the main concern of this section (inductive simplicity), let us briefly dispose of the possibility of satisfactory comparison based on some other uses of 'simplicity'.

It is sometimes claimed that we should try to arrieve at theories which are simpler, in the sense that our theories should describe simpler states of the Universe. We apply the predicate 'is simple' to the Universe according to attributions of certain characteristics to it. Can we use this concept of ontological simplicity in helping to decide between competing theories?..If we adopt a belief that structural properties of the world have the same _form_ as the propositions which make up the descriptional part of the theory (a 'picture' theory of meaning), then we might be able to make ontological simplicity claims about the world, on the basis of an examination of the form of these propositions. However, such a belief seems fundamentally obscure; it has been severely critisised.

Popper points out that "any idea of a particular structure of the world - unless indeed we think of it as a purely mathematical structure - already presupposes a universal theory."[12] Our ontological simplicity claims, in other words, are dependent on just what theories we hold. There is no prospect of using any concept of ontological simplicity in helping to decide between theories, for we have no option but to consider the world as complex or simple as our theories of it.

Now let us consider whether _notational simplicity_ can play any role in deciding between theories. We attribute notational simplicity to _descriptions_, i.e. it is a _linguistic_ notion as opposed to an ontological notion. We can distinguish two types of cases where we would claim that a description had notational simplicity: (1) where the attribution is independent of any person's psychological responses, (2) where it is so dependent. In the former case we would call attention to such aspects of the description as its brevity, or the number of typographical characters. In the latter, we would call attention to the familiarity of the notion, or its ease of manipulation, or its aesthetic appeal. In the former case we are dealing with _objective_ notational simplicity and in the latter case with _subjective_ notational simplicity. I shall discuss these saparately.

We can indeed conduct comparisons of type 1, where the rules of comparison determinedthe same objective notational property of the theories, say, brevity of the universal statements. But in the light of _what_ secondary goal would we choose the briefer theory? We could suggest, 'To arrive at theories whose universal statements are briefer' but this leaves us in the dark, as to why it would be beneficial to achieve this goal. It may be that, if pressed, we could give some satisfactory account in a particular case - we might, for example, be pressed for space in a text-book we are compiling. If however, we consider the goal, 'To arrive at briefer theories' to be applicable to <u>comparable</u> theories, then it seems difficult to give any justification for its achievement, unless we assume that objective notational simplicity in some way parallels the logical simplicity (q.v.) of the statement. Any attempted justification suffers from the critisism that it is always possible, by some suitable choice of vocabulary, to transalate any statement into one of minimal length. Unless this general goal operated under some severe constraints, we would arrive at one-word theories.

Of subjective notational simplicity, Popper has to say:

> "It is sometimes said of two expositions . . . that one is simpler or more elegant than the other. This is a distinction which has little interest from the point of view of the theory of knowledge; it does not fall within the province of logic, but merely indicates a preference of an aesthetic or pragmatic character . . . In all such cases the word 'simple' can easily be eliminated. Its use is extralogical."[13]

It is, of course, true that subjective ontological considerations play no role in a logical account of the growth of knowledge; whether one theory explains more or less than another has nothing to do with (say) the relative ease of manipulation of the equations of each of these theories. Nevertheless, scientists are not logical automatons with perfect powers of manipulation. Any historical account of the growth of knowledge will include references to preferences based on subjective notational simplicity. The poverty of mathematical progress in England in the eighteenth century, compared to the advances being made by mathematicians on the Continent, was, in large part, due to the difficulty of manipulations of the Newtonian Calculus of Infinitesimals in marked contrast to the ease of manipulation of Leibnitz's notation. It was no coincidence that it was the Continental mathematicians who advanced new theories based on Newtonian gravitational theory to account for planetry perturbations and the tides, (Bernoulli and Laplace) rather than their British counterparts.

Now, although Popper professes to be only interested in a logical account of the growth of knowledge, he seems willing to subscribe to such methodological principles as 'The aim of science is to make mistakes as quickly as possible - in order that we may learn from them.'[14] As soon as we adopt any methodological principle which concerns the rate of making mistakes, then questions concerning the ease of manipulation of the notation of theories become relevant. If we can more rapidly produce deductive consequences which provide severe tests of the theory, then we maximise our chances of refutation and hence of discovering new problems.

Aesthetic considerations must be distinguished from pragmatic ones. Whereas we can give grounds for preference in the latter case, e.g. T_1 can be used to make new predictions more rapidly than T_2, in the former wecannot. Earlier we eliminated such 'non-critical' appraisals. Even when we consider just comparisons based on betterness in respect of the achievement of some pragmatic goal, we are faced with enormous difficulties in assessment. If we attempt to measure 'degrees of notational simplicity' in terms of 'rate of production of new deductive consequences' we must make the assumption that the one <u>does</u> provide a measure of the other. It is by no means clear that this is generally the case, and to make any assessment <u>consistent</u> we would need some (obviously absurd) notion of a 'standars scientist' whose 'rate of production of new deductive consequences' could be interpreted as a measure of the notational simplicity of the theory, for , if we had no such notion, any assessment of notational simplicity would depend on the skill of the scientist of the day.

Although we can, for many reasons, change our relative assessment of two theories, we cannot make the grounds for such changes stem only from the effect of replacing one scientist with another. Although we <u>may</u> say 'The aim of science is to make mistakes as fast as possible' we <u>cannot</u> claim that 'making mistakes quickly' is a ground for objective choice between theories, and we can argue similarly for any other pragmatic goal which is relevant to subjective notational simplicity.

Now, let us turn to the main part of this section. Many philosophers have claimed that we can decide between theories on a basis of their logical or structural simplicity. It is this notion of simplicity which is used when Wittgenstein for example, says: "The procedure of induction consists in accepting as true the simplest law that can be reconciled with our experiences."[15] This use of 'simplest' and the comparisons associated with it form part of an inductivist approach to the growth of knowledge. Our theories are 'suggested' by certain observational data from which, if they form a regularly observed pattern or sequence, we induce a universal statement of a theory. If we have a situation where a set of (experimental) data is determined by experiment, the problem for the inductivist is how to choose between the various theories which have as one of their deductive consequences a statement, usually in the

form

form of a theory equation, which describes the observed experimental situation.

Several solutions to this have been proposed but I shall consider only one, that of Jeffreys,[16] which is, however, typical of the many suggested procedures. This solution consists in finding a simplicity ordering of <u>all</u> possible theories in terms of the number of parameters contained in the theory equation and directing the choice between the ones which are compatible with the data to the one which has the highest degree of simplicity. The problem is reduced to establishing a priority ordering between theory equations which can generate a curve which the experimental values 'fit' to some desired degree of accuracy.

We thus need to be able to determine just which theories are compatible with the data - how nearly does the situation described by the deduced statement have to tally with the experimental results to be 'reconciled with our experiences'?

Let us examine what is involved in the claim that we can use this <u>a priori</u> comparison of the logical simplicity of two theories to choose between them. Firstly, it implies that we have adopted the secondary goal 'To arrive at theories of greater logical simplicity'. Secondly, it implies that we have some set of rules by which we can order possible theories in terms of the number of parameters in the theory equation, and that this ordering corresponds to an ordering in terms of logical simplicity. Thirdly, it implies that some rules enable us to determine just which theories are compatible with the data.

Are there: (1) any satisfactory grounds for the adoption of this goal? (2) any reasons to believe that we can find such a set of rules which can be interpreted as giving an ordering in terms of simplicity? or (3) any reasons to believe that we can find a set of rules by which we can determine which theories are compatible with a given set of data? I will rely heavily on an argument put forward by Ackermann,[17] which suggests strongly that the ordering proposed by Jeffreys cannot satisfy the second implied requirements. But I shall go on to suggest that <u>any</u> method of ordering which attempts a solution of the problem in the same way, i.e. an <u>a priori</u> ordering of all possible theory equations, is inherently unsatisfactory.

Ackermann's argument is that Jeffreys' investigations have artificially made the problem too simple. Firstly, the proposed method of ordering takes account only of the number of parameters, (i.e. variables in the equation which, for a particular examination of the relation between other variables, is held constant). This method gives an intuitively reasonable method of ordering polynomials. It seems, however, that polynomials were the _only_ functions with which this simplicity ordering was expected to deal. It _does_ provide an intuitively acceptable ordering of this class of continuous functions, but there are other types of function which cannot be included in the ordering in an intuitively acceptable way. The transcendental functions $y = e^x + c$, and $y = \log x + c$, for example, would be ordered as _extremely_ simple. This ordering does not seem to correspond to any intuitive ideas we have as to the relative simplicity of transcendental and continuous polynomials.

The problem is even worse when we add _dis_continuous functions. Why should we regard a continuous function which has n parameters as of the same simplicity as a discontinuous function with n parameters? In this case, as it were, the two types of function bear their different logical structures on their sleeves, (or at least on the curves which they generate). There seems to be no justification at all in classifying them as to their degree of simplicity, with polynomials containing the same number of parameters.

These are points put forward by Ackermann, as to why he considers such orderings as Jeffreys' not to be'well formed'. We can point out another, though perhaps not so serious, defect in Jeffreys' account. Take, in the customary symbols, the general gas equation PV = RT. In a particular experimentato determine the dependence of pressure upon volume, we can make T, the temperature, constant, and so make T a parameter of the equation. We would produce a family of curves of P against V at constant T. This procedure, however, presupposes that we are plotting our results in a two-dimensional space. If instead, we plot our results in a three-dimensional space, then the idea of T being a parameter really becomes redundant. Instead of a family of curves, we would plot a mathematical surface, to where T does not have, in the sense that it had before, any constant values. If we resort to an n-dimensional mathematical space, then, for some suitable value of n, we can reduce the number of parameters in any equation to zero.

The idea that our data always represent a curve in two-dimensional space and that the 'fit' should be between points and curves, rather than curves and surfaces, or points and surfaces (of any dimension), seems another oversimplification. But as mentioned, perhaps this is not so serious, as we could stipulate conventionally, a 'reference space' i.e. twO-dimensional space which we require to be used for the plotting of data.

Ackermann has other arguments as to why the whole problem of arriving at accepted theories, by a curve-fitting technique, runs into problems. All experimental values must be measured in terms of rational numbers. How then, can we explain the presence in the body of scientific theories which need for their expression, real numbers?.."The occurence of and e, for example, cannot easily be explained on the basis of extrapolation from observed data".

Another problem, is that it is frequently assumed that we have only one set of experimental points through which to construct our theoretical curve. What is the scientist to so when experiments conducted on several occasions produce conflicting results? Which set of points is taken to be the one with which the theoretical curve must be compatible? Thus, not only nust we solve the problem of deciding which theories 'fit' the observational points, but also we must decide just which experimental points shall be the ones to which we try to match the curves.

If these assumptions are brought out into the open, then it seems that none of them have any justification. If they are dropped, then no procedure so far suggested, including Jeffreys', seems adequate to meet the implied requirements. This does not mean that some means of ordering all possible theory equations <u>might</u> not be possible, but any solution to this would have to be much more powerful than any suggested so far.

But even given such a well-ordering, it remains to justify this ordering in terms of some explicit notion of simplicity. If all that were required were an ordered list of all possible hypotheses through which the scientist worked until he came to one which was compatible with his data, then the list could be ordered in terms of any notion we pleased. In other words, it may be possible to arrive at some set of rules, by which we can say of any two theory equations, "T' is better than T because it is better in respect x", i.e. that the rules of comparison are such that we achieve the primary goal, but unless we can justify the claim that 'better in respect x' does in fact correspond to 'simpler' (some <u>explicit</u> notion of simplicity), we cannot claim that the rules compare the respective simplicity of the theory equations. But even if we can justify this claim, we have to explain why we should try to achieve the secondary goal 'To arrive at theories which are (in this explicit sense of the notion) simpler. For reasons which we have given above, we can rules out any possible justification based on the idea that the adoption of simple theories has some pragmatic benefit. One popular idea is that simpler theories, (in the sense of fewer parameters) are more probably true, but this seems to involve the idea that Nature in some way, prefers to do things in a simple manner. Such an idea as this seems fundamentally obscure, and relies on notions of ontological simplicity which we previously dismissed.

Weyl,[18] in discussing a situation which has close parallels with the curve-fitting problem discussed above, also argues against the applicability of probability to justify our acceptance of the simplest curve. Consider the case where we have a set of observational points which lie very nearly on a straight line, but any curve which passes <u>through</u> those points, must be a polynomial of high degree. Weyl suggests, but dismisses the idea that we are justified in accepting the simpler curve (the straight line) because it is extremely improbable that the experimental values, whose position on the curve is quite arbitrary, should fall very nearly on a straight line if, in fact, the function were not a straight line function. This, Weyl points out, will not do. We can draw an infinite number of given points, and for <u>every</u> curve, it would be true to say that it would be extremely improbable for the points to lie on this curve if, in fact, the law were not represented by this curve. By this argument, all these curves would be equally probable, and yet many would deviate from a straight line.

My conclusion from this discussion on inductive simplicity and the curve-fitting problem, is that (1) no satisfactory ordering of theory equations has yet been presented, and that (2) the notions of simplicity involved in the orderings is so vague that the desirability of the adoption of any secondary goal based on them is questionable. Now, I wish to go on to question whether the whole approach of <u>such a priori</u> comparisons of theories is, in <u>principle</u>, satisfactory.

We must look more closely at what is expected of this type of competition. Let us assume that we have some method of ordering theory equations which copes satisfactorily with all the types of function we have considered. If we assign the same degree of simplicity to more than one equation, we would arrive at an ordering of subclasses of the class of functions which had the same degree of simplicity. If our method of ordering was sufficiently discriminating, these subclasses might have only one member. Our experimental results in hand, we now work down this ordering until we reach a theory equation which is compatible with the data. (We will assume that we have some method of assessing compatibility.) _This_ is the simplest theory compatible with our data, and hence it represents the general law, and we might add some such comment as 'it is the _most probable_ theory'. What is expected of this _a priori_ assessment is that if our ordering contains _all possible_ theory equations, then, sooner or later we shall come across a theory equation which _is_ compatible with the data. (I shall neglect here the problem that we might have to eliminate an infinite number of equations (say) in the simplest subclass before we could move on down the list.) But let us examine just what is meant by 'all possible theory equations'. It can surely only mean 'all possible theory equations which are expressible in the most powerful mathematical languages that we have'. We could hardly insist that the list contained theory equations which needed for their expression, types of functions not yet discovered (which needed for their expression a language more powerful than any yet developed), for an inspection of the list would bring them to light. In any case, it is doubtful whether we would have any justification in claiming that our ordering method satisfactorily ordered unknown functions (which might have quite peculiar simplicity properties).

So mu first objection is that it is possible in any given ordering, that we shall <u>never</u> come across a theory equation which is compatible with the data, simply because the theory equation which describes these data has not been included in the ordering. It might be objected to this, that for <u>any</u> given set of points, it is possible to construct <u>some</u> curve which passes exactly through them and hence we <u>must</u> sooner or later, even if the ordering was just of polynomials, come across the simplest theory. But, this objection is easily countered. We know that a present-day satisfactory ordering of theory equations, would include discontinuous functions. This implies that if the procedure is to work <u>at all</u>, there must be some method of discriminating discontinuous from continuous '<u>data-plots</u>' (This could be done, perhaps, by examining small regions around each datum point to test for continuity of the observational variable.) So we could rule out the possibility of compatibility of the data with <u>any</u> polynomial.

What conclusions can we make in the light of this? Those who claim that if (in principle) this procedure was followed, we would inevitably end up with the most probable theory, must refute the charge that, in principle, there are cases where no theory could be found at all. It is surely unreasonable to assume that all laws which have as yet not been formulated _could_ be so formulated using only the mathematical functions we have at our disposal at the present. My argument suggests that there are conceivable cases where, in principle, the inductive procedure breaks down. It breaks down, essentially, because the comparative procedure assumes that there is _a_ theoretical language in which it is possible to express all possible theories, but I do not believe that this is the case. Our rules of comparison, by which we rate the theories as to their degree of simplicity, are dependent upon a particular language and if that language changes, we have no guarentee that the old rules will be adequate.

There seems no reason, in principle, why the method of ordering which is satisfactory for n types of function, would be satisfactory for n + 1 types of function, and moreover, there seems no reason why a method of ordering which supercedes the old one should necessarily retain the original simplicity ordering for the old class of n types. We _could_ stipulate that the rules which determine the ordering may _only_ change in such a way as to preserve the original simplicity orderings unaltered, but whether this is a practical possibility, or an unattainable ideal, is problematic. (A similar stipulation concerning the restriction of possible modification of confirmation functions can be shown to be unattainable. This is (roughly) that modifications shall be such that evidential support assessed in an old language is not revoked by transition to a new. (We shall examine this in the next section.))

If it is the case that revoking of simplicity orderings does occur on the introduction of a new language we must face the possibility that the curve-fitting procedure results in a different choice from that made in the old. <u>The</u> simplest curve now becomes meaningless. We can speak only of 'the simplest theory equation relative to language L'.

Let us consider this result in the light of the inductivists' original programme. In part, this was to assign degrees of simplicity to theory equations; but this assignation was the result of application of a procedure of a purely formal character. We could regard the comparison rule which governs the assignation of the degree of simplicity as (in a suitably constructed logical calculus), a rule of inference. From a statement describing the form of the theory equation, we infer with <u>certainty</u> a statement describing the degree of simplicity of the equation. But, insofar as this is a formal rule of inference, all this amounts to is a logical truism similar to '2 + 2 = 4'. The statement 'If theory x has n parameters (or whatever), then theory x has a degree of simplicity m', does not state any fact about theories any more than '2 cats + 2 cats = 4 cats', states any fact about cats.

It may be objected that in a sense '2 cats + 2 cats = 4 cats' does state a fact about cats; it enables us to calculate what happens when, to two cats, we add two more: viz the cats do not disappear or polymerize, but remain 'discrete' cats. Now, however, we are <u>interpreting</u> '2 + 2 = 4' as a physical theory rather tham a logical one, and it becomes open to falsification. It <u>would</u>, in fact, be falsified by '1 drop of water + 1 drop of water = 1 drop of water'.

We can interpret the rule of inference mentioned above in a similar way. This would be an interpretation quite foreign to the proposers of the inductive curve-fitting strategem, for they would have to abandon their implied claim that there can be a purely logical measure of simplicity. But we have seen in the example given that we can conceive of cases where a change in theoretical language could lead to a revision of what we consider to be the simplest theory, and if a new theory in this new language stood up to severe empirical tests we would certainly consider the new language to be better than the old. This, of course, would lead to a <u>change</u> in the rule of inference by which we assessed simplicity.

This then is the objection to such a formal procedure. If they are interpreted as <u>purely</u> formal, their assessments of simplicity are logical truisms. If these assessments are interpreted as synthetic statements then they areadependent (in principle) upon just which language is used to express the theoretical statements of science; but the choice of language is essentially an empirical matter: we choose a language if the theories expressed in that language are empirically successful.

There are, it is true, escape routes from this state of affairs. (1) We could produce a language which could be shown to be capable of expressing <u>all</u> scientific theories. (2) We could produce a rule of assigning degrees of simplicity which we could demonstrate left original simplicity values unchanged on its modification to suit (any) new language. (3) We could avoid language dependency by introducing a new higher-order simplicity-assessing-rule which took into account the particular language used. Instead of arguing from (say) number of parameters to degree of simplicity, we would argue from degrees of simplicity in language L to (higher order) simplicity. We have, however, no reason to believe that (1) could be achieved. It is difficult to see what would count as a satisfactory demonstration. (2) seems to be equally problematic, and (3) leads to an infinite regress, (the higher-order simplicities are relative to another (higher-order) language which might be revised for similar empirical reasons).

This, then, is my argument for declaring the comparison of theories by purely formal methods of assigning degrees of simplicity to be unsatisfactory. Such formal methods claim that they can arrive at a result by a <u>certain</u> analytic procedure. In so far as this is so, the results are simply tautologous statements. But the proponents of these methods also claim that their analysis enables us to arrive at synthetic statements which describe actual properties of theories. If <u>this</u> is so, then we can show that the envisaged procedure is no longer analytically certain, but must be construed itself as a fallible theory about theories.

<u>SECTION B</u>

In this section I will deal with only a tiny part of the enormous field of confirmation functions. I wish to amass more evidence that the notion of purely formal appraisal of theories, (in the sense that the appraisal is analytically certain) cannot perform the function which is asked of it by its proponents.

As was stated in the Introduction, the attempt to assign probabilistic support functions to theories was one method of avoiding the result that our lack of ability to prove theories to be true indicated they did not constitute knowledge. The probabilists' programme was to develop an inductive logic which would enable the precise degree of probability of a theory to be calculated with respect to the available experimental evidence.

Carnap's original programme which was to assess the degree of evidential support for theories came to grief when he realised that intuitive concept of confirmation, which he thought could be measured by p(h,e), did not correspond to degree of evidential support which Popper had shown not to be probabilities. After a brief sojurn with 'qualified instance confirmation' which, he realised, gave a measure of the <u>reliability</u> rather than the probability of theories which he was looking for, he turned to his theory of 'rational betting quotients' and 'degrees of rational belief'. It is this which I wish to examine.

The first point to note is that it is only by courtesy that we consider it to be a method of appraising <u>theories</u>

at all. It is essentially an atheoretic system. Carnap determined to interpret confirmation as a probabilistic support function and sought some intuitive concept which could satisfy this requirement. He believed he had found this in the concept of a rational betting quotient, which measures, relative to the available evidence, a value which we are prepared to bet with good reason on the prediction of a single event. He strengthened his claim that rationsl betting quotients were probabilities by support from the Ramsey de Finneti Theorem. One condition of rational behaviour of a person who undertakes a bet is that the odds are not stacked against him - that he shall not certainly suffer a net loss when the final tally of wins and losses is made. If we call, with Carnap, a system of degrees of belief for a given field of propositions, a credence function, then a credence function is <u>coherent</u> if it excludes bets where loss is unavoidable. Now, as the Ramsey de Finneti Theorem proves that a credence function is coherent if and only if it satisfies the calculus of probability, Carnap felt justified in taking his rational betting quotients to be probabilites. Although this added support to the claim that this confirmation function is probabilistic, it left Carnap in a dilemma. His new calculus had produced the result that the confirmation of <u>any</u> universal statement was zero, $c(u) = 0$, it is a lemma of the Ramsey de Finneti Theorem that the probability of any contingent proposition is not zero ($p(u) \neq 0$). The result was that he dropped universal statements (theories) from his probabilistic appraisals; the rational betting quotient was to be independent of any theory. This makes our appraisals of theories derivitive upon our appraisals of particular predictions.

This result in itself causes severe criticism to be directed against the theory which produced it. We surely <u>do</u> take theories into account when deciding whether to bet on particular predictions. If we take the extreme case, we could consider betting on a <u>new</u> deductive consequence of a theory which had been well corroborated in other fields. Carnap's theory would give a very low rational betting quotient for this as no evidence would as yet be available. However, the scientist would be prepared to bet more heavily <u>because</u> it was a deductive consequence of an already well established theory.

But even within its own terms the programme can be shown unsatisfactory. The rational betting quotient measures our degree of rational belief in a particular hypothesis and we arrive at these values using an analytic inductive logic. Carnap, wishing to avoid any problems of the dependence of degrees of confirmation upon the progressively expanding theoretical language, stipulated that these should be measured relative to the 'minimal language' in which the particular predictions could be expressed, thus any change in language would leave the previously established confirmations unaltered. But, as Lakatos has shown, even if we accept that this is feasible, we <u>still</u> run into problems of language dependency because of the possibility of indirect evidence. "Indirect evidence relative to L in L* (would be) an event which does not raise the probability of another event when they are both described in L, but does so if they are expressed in a language L*". An example of this: in the language of Gallilean mechanics the motion of the satellite of a planet would not be regarded as evidence which made us assign a higher confirmation to the law describing the trajectory of a projectile, whereas if we express these events in the language of Newtonian mechanics, it would.

It seems therefore, that a change in language can produce the result which Carnap wished to avoid - the change of confirmation values of previously evaluated particular predictions; once again the hopes for the procedure cannot be realised.

I have included this brief and wholly derivative sketch of the argument against the satisfactoriness of using analytically determined confirmation functions, merely to illustrate what is expected of a certain type of procedure and some criticisms which can be directed against it. It shows some similarities with the previously discussed inductive-simplicity method. Of each programme there is much more that could be said in its defence, but I have introduced them, not mainly for the purpose of direct criticism, but to act as a background for discussion of some falsificationist comparisons in the next chapter. Nevertheless, the criticism outlined is powerful. If we compare theories either directly of indirectly by following rules of comparison which we claim are _analytic_, then we cannot, in general claim that our appraisals are anything more than tautologies. If we _do_ claim that they have synthetic content, then they become falsifiable, and hence not certain. If we break the circle by claiming that our appraisals are _a priori_ synthetic statements (in a Kantian sense) then we are faced with the same problem on a higher level: either we are faced with an infinite regress of appraisals, or we retreat to _a priorism_.

CHAPTER THREE

The notions of betterness used in the falsificationist methodology which concerns us here are closely related. The falsificationist's claim is to have shown that all the main epistemological problems stand in a systematic relation to one another, and moreover, that these problems can be solved in terms of testability, falsification and corroboration, and total corroboration.

Against this background, and with the results so far obtained I wish to tackle the following problems. (1) How can we demonstrate the need for representatives of the three methods of comparison listed earlier? (2) On what grounds can we avoid the charge that the type of criticism outlined in Chapter Two are not applicable to these falsificationist appraisals. It is my contention that these questions are in fact quite closely linked.

Now, there is a 'fast' answer to question 2, which stems from the whole concept of this theory of method. This is to look upon the whole of scientific activity in much the same way as we would look upon (say) a game of chess. It is a game where certain moves, the acceptance or rejection of theories, are 'permitted' by methodological rules which determine how this game is to be played. In order to decide whether a proposed move falls within the rules or not, we conduct a certain analysis of the logical relations which hold between statements. If the rules are satisfied, then on the basis of this satisfaction, we can make an appraisal but this appraisal is nothing more than a <u>tautology</u>. All that it claimed is that, for some particular notion of betterness, this notion is <u>defined</u> in terms of the logical relations which hold between the analysed statements.

Let us consider the falsificationist appraisal 'Theory T has a higher degree of corroboration than theory T'.' This appraisal can be made if and only if certain relations hold between certain statements, viz: there is an accepted basic statement which is prohibited by theory T, but is not prohibited by theory T' and there is another basic statement which is a deductive consequence of T' but not of T, such that this basic statement was not, prior to the introduction of T', accepted, but after empirical testing of T', was. If these relations hold then, <u>by definition</u>, T' is more highly corroborated than T. (This is an oversimplified list of the relations which must hold, which I am using just to illustrate a point.) We make no claims that this appraisal is synthetic, in the sense that it gives any more information about the theory than that these relations hold. In particular it makes no claim concerning any future performance of the theory; hence it is obvious that the criticisms which applied to the appraisals of Chapter Two are not applicable here.

This interpretation of the methodological rules which govern the game of science is open to a very grave objection. Although it is reasonable to apply this sort of analysis to the game of chess where the players accept that certain rules shall be followed 'for the sake of the game' we may ask why does (or should) the scientist agree to play the game of science by following these rules. If the scientist's appraisals are tautologies, then the end result of his activity will be a body of statements compiled according to a set of rules, but nothing more. He will have no reason to believe that this abstract procedure produces a body of statements on which he can rationally base any practical action, and no reason to believe that superceeding theories are closer to the truth. And yet we expect that the scientific enterprise will do just this - we expect that our most recent scientific theories may, with only slight modification, be equated with the body of technological theories upon which we can rationally base our actions and we expect that our superceeding theories will be nearer the truth. If this is the case, then we must be able to offer some explanation as to why an adherence to the proposed methodological rules will produce such a result. But can this be done without once again falling victim to the old criticisms ? For any interpretation of an appraisal as a claim to be assessing the future reliability of theories surely must involve a synthetic statement. As these are in principle falsifiable, as we have seen, how can we justify this appraisal without retreat to a priorism or infinite regress ?

We can overcome these problems by introducing a concurrent secondary interpretation of at least one of the types of falsificationist appraisal. This interpretation is, however, dependent upon a metaphysical speculation which we catagorically assert to be fallible. Thusnwe establish a duality of appraisals; the first, to which the notion of fallibility does not apply, is that appraisals are tautological, and the second is that they are speculative synthetic assertions which we freely admit are fallible.

Both of these interpretations escape the criticisms applicable to inductive appraisals. In this non-justificationist theory of method the idea that our appraisals must be justifiable (i.e. shown to be true, or probable to a certain degree) is absurd. We must, though, be able to give some fairly convincing reasons why the tautological appraisals can be given such a secondary interpretation as to enable us to consider the most recent theories of science nearer the truth and more reliable than their predecessors. If we can do this successfully, it will go some way to support the claim made on p. 38 that behind the adoption of secondary goals of science is the implied claim that this will lead to epistemological or technological advance.

Let us look at the falsificationist notions of: (1) testability, (2) corroboration and falsification, and (3) total corroboration in turn and at the secondary interpretations which have been suggested for the last two groups of appraisal,(2) and (3).

(1) Testability.

Popper gives two sets of rules for comparing theories with respect to their degrees of testability (which he equates with "degree of falsifiability"). When both are applicable to the same competition they result in the same preferential choice being made. The first of these, which Popper describes as the 'more sensitive' gives a method of comparison by examination of classes of potential falsifiers. The class of potential falsifiers of any theory is, of course, infinite, and as, moreover, they all have the same cardinal number any comparison of the two theories by assignation of numbers of potential falsifiers breaks down. In cases where the classes stand stand in a (proper) subclass relation, however, we can state that one theory has <u>excess</u> testability. Popper defines 'empirical content of a theory' $=_{df.}$ 'the class of potential falsifiers of that theory', but it is highly dubious whether this absolute empirical content measures anything. It is also problematic whether we can assign to any theory in isolation a particular degree of testability.

What is more certain is that if the class of potential falsifiers of one theory, T, is a proper subclass of the class of potential falsifiers of another theory, T', then T' has a higher empirical content, in the sense that T', to avoid refutation, would have to be compatible with a greater number of observation statements than T. The subclass relation will only hold when one class is contained in the other, so in general, most theories can not be compared with respect to degree of testability by these rules of comparison. Among those theories which _can_ be compared by this method will be those rival theories which I have called _comparable_. It is because of this that this comparison has methodological value, enabling us to choose, _a priori_ from among several theories which may superceed an established theory, those theories which are most worthy of testing.

The method of the subclass relation cannot give any absolute measure for the testability of theories but the second method. that of comparison by dimensional number can. This method, which I described earlier, compares theories with respect to their _dimension_, the minimum number of 'relatively'atomic statements' needed to refute the theory, relative to a particular field of application. To achieve Popper's aim of establishing some measure of the absolute empirical content of a theory we would have to resort to this method. It is however, questionable whether the two sets of rules measure the same property. Although it is the case that, when both rules of comparison give grounds for the same preference, it does not follow that these grounds are the same. If the scientist asks why he should choose more highly testable theories, there seems to be two separate answers depending on which rules of comparison were followed. If the subclass relation, then we can say 'choose the more highly testable theory, because there is more available material against which to test the theory'. If the dimensional comparison, we would say '. . . because you will need a fewer number of observations to see whether it is refuted'. (Let us differentiate the two sets of comparison rules. If theory T, is rated as more highly testable by the subclass rule, call this more highly testable$_s$, if by the dimension rule, testable$_d$.) Although if a theory is more highly testable$_s$ it does seem reasonable to claim it has greater empirical content, it is questionable whether we can say the same of a theory which is more highly testable$_d$. We cannot argue that because both comparisons produce the same results they must measure the same quantity, which is what Popper appears to do. If T' is more highly testable$_s$ than T, then T' restricts to a greater extent the number of possible states of affairs which the theory sanctions, and in this sense it may be said to have greater empirical content. It is difficult to see how the same could be said of the second comparison. If T' is more highly testable$_d$ than T, then unless we can conceive of certain states of nature being

per se

<u>per se</u> more 'restricted', we do not seem to be speaking of the same thing. If this is so then we have no method for assessing absolute content which is intuitively satisfactory. Of course empirical content is <u>defined</u> by Popper in terms of testability and to that extent we must accept that this is the way he chooses to use the words. However, it is not necessary to assess absolute empirical content to carry through his methodological theory, <u>excess</u> content is sufficient, and we have, in the subclass relation a method of assessing (though not to any precise degree) this quantity.

In "The Logic of Scientific Discovery", Popper equates the notion of degree of testability with that of simplicity. He claims that this move results in us being able to answer all the epistemological questions which arise in connection with the concept of simplicity. However, in his later works he claims that only one 'important ingredient' in the idea of simplicity can be logically analysed in terms of testability. His later position leads to a 'Duhemian' type of appraisal where one theory is said to be better if it unites previously unrelated areas of investigation.
If the classes of potential falsifiers of two theories T_1 and T_2 were both proper subclasses of a class of potential

falsifiers of a third theory T_3, but did not themselves have any common elements, this would be grounds for claiming that T_3 unites the previously separated fields covered by T_1 and T_2.

We have claimed that comparison with respect to degree of testability provides a method for assessing which theory is more worthy of consideration. This is true if we have no, or very little background information; in general the more highly testable theory will be easier to refute. But consider the following (hypothetical) situation. Say some perturbation of Pluto's orbit is observed. A new 'ultra telescope' technique finds many tiny particles whose sum gravitational field accounts for the remaining one-quarter of the effect. (1) T_1 which predicts that the remaining perturbation is due to a single perturbing body in a particular elliptical orbit. (2) T_2 which predicts that the remaining perturbation is due to more and finer particles, but in many possible configurations. If we choose between these theories by comparison of their testability only, we would choose T_1 as this can be refuted in a minimum of six observations. But in the light of our background knowledge we would choose T_2, even though it would require more than six observations to refute it.

We would expect that the theory which had successfully explained three-quarters of the perturbation would (in a slightly modified form) hold good for the remaining one-quarter. It may seem as though some inductive principle has crept in here, that our choice is made on the assumption that an older theory will continue to hold good, but this is not necessarily the case. We could regard the proposal 'the remaining perturbation is due to finer particles' just as a proposal to test (although not severely) a theory which has not yet been falsified.

These _a priori_ appraisals are all tautological, in the sense that they are defined in terms of the relations which hold between statements. Moreover, they have not been given any secondary interpretation, so appraisals of competing theories on these grounds could not answer the questions we asked earlier concerning truth and reliability. It is also difficult to see how any secondary interpretations of these appraisals _could_ give any (even fallible) estimate of whether the theory preferred was closer to the truth or more reliable. If we demand (1) that the theory of method should contain this type of _a priori_ comparison and (2) that the theories must be able to be compared with respect to truth content and reliability, then these _a priori_ comparisons can only be one of a (perhaps related) set of comparisons. They do however, have a vital role to play in their own right, (1) in determining the potential test worthiness of theories, and (2) in determining the degree to which a theory is corroborated (q.v.).

(ii) Falsification and Corroboration

What interpretations can be given to falsificationist appraisals which enable us to say that superdeeding theories are nearer the truth? Firstly let us consider the tautologous appraisals, (1) 'Theory T is false but theory T' is not', and (2) 'Theory T' is more highly corroborated than theory T.' We shall consider the case where T' has been chosen a priori as of greater empirical content (by the subclass comparison). For (1) to hold there must be at least one potential falsifier of T which is accepted, but no potential falsifier of T'accepted. As the class of potential falsifiers of T is contained in the class of potential falsifiers of T', then T' must be consistent with the accepted potential falsifier of T. We might interpret the result of this competition as an appraisal of growth, but we could not on these grounds alone claim that this was empirical growth, as T' could have been arrived at by an ad hoc manoevre to ensure consistency with this accepted potential falsifer.

If however, the appraisal (2) holds then we can interpret this as an appraisal of <u>empirical</u> growth. For (2) to hold T' must have successfully overcome some severe, <u>independent</u> test, the predicted outcome of which was held to be extremely unlikely or even impossible with respect to the old theory T. We cannot assign any <u>numerical</u> degrees to measure how much a theory is corroborated, But we can to some extent assess excess corroboration of one theory over another. If a theory which is more highly testable is corroborated, in a more severe test, we could claim it had excess corroboration over a less highly testable theory corroborated in a less severe test. Testability can be assessed by an <u>a priori</u> comparison.

We can <u>define</u> empirical growth in these terms, moreover, such a definition seems intuitively acceptable. We can use the notion of higher corroboration in a secondary interpretation to give an appraisal of competing theories in terms of truth content. This interpretation relies on the notion of an adequate concept of objective truth. (I will accept Popper's claim that the Tarskian formal analysis can provide the basis for such a concept of truth which is applicable to the languages of sci-ence.) Now we can make the (admittedly) fallible claim that if a theory T' is more highly corroborated than theory T, then T' is nearer to the truth than T.

Popper defends this interpretation by giving a more precise meaning to 'nearer to the truth' which removes any idea that T', being nearer to the truth, is more <u>probable</u>. Any theory has, independently of its truth value, some degree of <u>verisimilitude</u>, which we can define in terms of truth and of empirical content.

"Let us consider the content of a statement A: that is, the class of all the logical consequences of A. If A is true, then the class can consist only of true statements.... ... But if A is false then its content will always consist of both true and false conclusions. Thus whether a statement is true of false there may be more or less truth in what it says, according to whether its content consists of a greater or lesser number of true statements. "

"Let us call the class of the true logical consequences of A the 'truth-content of A' ...; and let us call the class of the false consequences of A the 'falsity content of A' ...; Now we can say, assuming that the truth content and the falsity content of two theories T_1 and T_2 are comparable, ... T_2 is more closely similar to the truth, or corresponds better to the facts, than T_1 if and only if either:
(a) the truth-content but not the falsity-content of T_2 exceeds that of T_1, or
(b) the falsity-content of T_1, but not its truth-content exceeds that of T_2." 19

This interpretation of theories being closer to the truth, or having a higher verisimilitude, depends upon the fallible assumptions (1) the more highly corroborated theory has a higher truth-content, and (2) if a theory is falsified, then it has a higher falsity-content than a theory which is not falsified by the same potential falsifiers. In any particular case we can only guess that these assumptions are true. If asked 'how do you know that theory T is closer to the truth than T''? we can only reply that it is a claim to conjectural knowledge only.

It is apparent that the two separate conditions of the definition of verisimilitude can be satisfied by two separate appraisals. As a result of a comparison of the degree of corroboration of two theories we can speculate that one theory has a higher truth-content than the other. As a result of comparing whether a potential falsifier refutes one theory but not another, we can speculate that one theory has a higher falsity-content. To assess whether one theory has a higher degree of verisimilitude than another, we have to take both these appraisals into account. If we define the degree of verisimilitude of a theory as the difference between the truth and falsity-contents, then we can imagine that these degrees of verisimilitude correspond to locations in some matrix. It would be misleading to suggest that we could assign numerical values to the degrees of verisimilitude and hence set up some partial ordering relation for theories. But we can estimate <u>differences</u> in degrees of verisimilitude in terms of corroboration and falsification, and hence for two theories conclude that one is nearer the truth than the other.

If we are motivated by the goal 'to arrive at theories with higher verisimilitude', then logically it is immaterial whether we do this by (1) choosing theories with a higher truth-content, or by (2) choosing theories with a lower falsity content, or (3) both. From a methodological point of view, however, these three methods are not equivalent. Method (2) could be achieved purely by a content decreasing <u>ad hoc</u> stratagem, i.e. by a modification of a theory so as to make it consistent with an accepted falsifier. Are methods (1) and (3) methodologically equivalent ? I think that they are not. This point can be illustrated by considering two sequences of theories. One,(sequence A), is a sequence where each succeeding theory has a higher truth content, but the same falsity content. In the other,(sequence B), each succeeding theory has both a higher truth-content and a lower falsity-content than its predecessor. Sequence A represents a progression of theories where each member of the series successfully predicts new facts not predicted before in addition to a basic statement which refuted the previous theory. Sequence B preceeds by refutation and c corroboration, each stage coming into <u>conflict</u> with experimental evidence. Because in each succeeding theory, we learn from the mistakes made previously, we can guess that this sequence represents genuine empirical pregress. In sequence A, however, we make no mistakes - as this sequence progressed we would become increasingly suspicious that the theories were loosing their empirical character. Even though our predictions were successful we would realise that the foundation on which we base our claim that each succeeding theory is nearer to the truth was becoming more remote, for as each succeeding theory is not framed in the light of some <u>problem</u> which arises from the anomalies which faced the old theory, the choice of new theories becomes increasingly arbitrary. We could conceive of a bifurcation of the sequence, ateach stage two apparently compatable theories could be introduced but with no experimental conflict we would be able to choose between them only on the grounds that one was more highly corroborated . But the degree of corroboration depends upon the degree of testability of the new theory which is interpreted as its degree of falsifiability. With no falsification at all, these concepts all become rather empty.

It may be objected that it is unreasonable to deny that sequence A represents a progression of a series of theories which have increasing verisimilitude. I would agree with this. My point is that from a methodological point of view this notion is <u>derivative</u> upon the notion of approaching the truth by simultaneous increase of truth-content and decrease of falsity-content. If we choose a more highly corroborated theory over an unfalsified one, we expect that <u>if</u> we conducted the relevant experiments, we <u>would</u> be able to demonstrate inadequacies in the old theory in those areas where the new theory had higher content. Of course, cases <u>may</u> arise where this does not occur. Are we to say of such cases that we do not consider the acceptance of the new theory methodologically sound ? It would be nonsense to suppose that many such theories did not represent a genuine advance in truth-content but, nevertheless, in any sequence of superceeding theories we must severely restrict the occurence of acceptances of this type if we are to ensure that the empirical growth remains on a firm empirical basis. If the sequence deviates greatly from the schema 'conjecture, corroboration and refutation', it becomes increasingly unsatisfactory. To prevent this degeneration from the empirical basis, therefore, it is essential that our comparison methods include type 2 comparisons as a result of which we can claim that the previous theory was false.

We must now consider whether this ascription of falsity applies to individual explanatory theories, or to a whole theory system. So far in this chapter, I have not made this distinction. In the introduction I characterised naive methodological falsificationism (NMF) as a theorybof method which included 'conventional' falsification of theories and called this 'falsification$_P$'. Does Popper's methodology include such falsifications ? Several times he explicitly states that we falsify only the total theory system, but I wish to argue that it is more in keeping with the spirit of his theory to consider individual explanatory theories as being conventionally falsified. Although he cannot be unequivocally classed as either a naive or a sophisticated falsificationist, I believe his position is closer to NMF than to SMF.

Although fallible conventional decisions are needed to falsify$_P$ a theory, I shall argue that the conventional element is not as great as has been suggested by, for instance, Lakatos and that it is advantageous to take these decisions. (In Chapter Four I shall point out certain changes inherent in not taking them.)

The conventional decisions needed for refutation of a theory are those that isolate the theory from the 'unproblematic background knowledge'. What grounds can we give for claiming that when an experimental result indicates an inconsistency between accepted basic statements and the deductive consequences of the theory, we may interpret this as refuting the theory?.

Firstly, we may note that the view that each empirical test we make puts on trial the whole of the body of empirical knowledge (one interpretation of the 'Duhem-Quine Thesis'), it is open to criticism. Although, in one sense, it may be trivially true, this does not mean that we cannot in many cases isolate that part which is responsible for the refutation. It is trivially true in the sense that we can never be certain that we isolate the right part, but several arguments can be put forward which suggest that our conjecture can withstand criticism.

Popper gives one based on independence proofs of axiomatized systems. Suppose we can show that a certain axiom of an axiomatic system is independent of all the other axioms, i.e. cannot be derived from them and that this axiom system can be given such an interpretation which allows us to predict that certain things do not happen. "There is no reason whatever why (a particular) counter example may not be found to satisfy most or even all our axioms except one whose independence would thus be established." This shows how we can at least in some circumstances, isolate a particular part of the axiom system as being responsible for the erroneous prediction. But it is not enough merely to demonstrate how a particular axiom of a system might be isolated. We need to justify our decision to isolate the theory from the total theory system which is relevant to a test situation.

In this example, Popper takes the prediction to be a non-existential singular statement. As these follow logically from the universal statement of the explanatory theory we do not need any statements of 'initial conditions'. The only recourse to an observational theory needed is in the decision to accept the basic statement which was considered a counter example of the theory. However, to deduce singular existential statements from a theory we need statements of initial conditions. Here we would need a second recourse to an observational theory, in the light of which we make the decision to accept these statements of initial conditions.

Now let us try to meet the objection that falsification_P of a theory needs the risky decision to accept the observational theories of the theory system as unproblematic. Are these decisions so risky? Firstly let us distinguish between the observational theories needed to accept statements of initial conditions ($O.T._1$) and the observational theories needed to accept basic statements which 'refute the theory' ($O.T._2$). As an example consider our 'theory', 'All swans are white'. To deduce the observation statement 'There is a white swan at x,y,z,t,' we need the initial condition 'At x,y,z,t, there is a swan'. We accept this statement in the light of a set of OT_1s which would include some theory of coordinate measurement and some ornithological theory of bird recognition in the light of which we claim that this particular bird is a swan. Now consider the OT_2 needed for the acceptance of the potential falsifier 'At x,y,z,t there is a black swas'. Obviously we will find in the class of OT_2s all the members of the class of

OT_1s, but in addition some 'colour recognition theory' such as 'All surfaces which absorb a certain proportion of incident light are black'. (In general the class of OT_1s is a proper subset of the class of OT_2s, and for this reason we cannot claim that refutation of a non-existential statement is a 'less risky' method of falsifying a theory.)

Thus, in general, if we succeed in justifying our acceptance of the OT_2s we will have also covered the OT_1s. What argument can be given in support of this acceptance? Firstly we point out that observational theories accur in a large number of different theory systems. Observational theories needed for measurement of mass, length and time are so ubiquitous and occur in so many theory systems which have been well corroborated that it seems perverse to suggest that in a particular system they could be responsible for the inconsistency, providing that the magnitudes of the quantities measured are such as to fall within the range of magnitudes previously dealt with by the observational theories in the well corroborated system.

Now if, for each observational theory, we can find an occurence in a previously <u>corroborated</u> theory system, we can reasonably conjecture that in the theory system under test we can attribute the origin of inconsistency to some other component than the observational theories. It might be objected that this condition will only rarely be achieved but this is not the case. As fields of scientific inquiry become more developed they also become increasingly unified. At the same time the obsevational theories needed become increasingly fundamental and consequently more highly interchangeable in theory systems. Such an argument is, of course, by no means a proof that this is so, but we do not claim that it is.

Secondly, we can point out that an inconsistent theory system TS can be restored to consistency TS' by substituting one explanatory theory T by another, T'. If both TS and TS' contain the same observational theories then we have good reason to believe that the previous inconsistency was due to the theory T, and not to the observational theories. It may of course, be the case that both T' and the observational theories are false and that T was not, the resolution of the inconsistency being due to the introduction of the false theory T'. If we also stipulate that TS' should be corroborated in a new test, we can conjecture that the likelihood of this is very small. Any criticism of a theory, if it is not gdng to involve an appraisal of the whole body of scientific knowledge relevant to that theory, must take some background knowledge for granted, or criticism itself would be impossible, (or at least an extremely lengthy procedure). We claim that it is reasonable to assume that such observational theories are unproblematic.

The decision to accept the ceteris paribus clause is also needed to falsify$_p$. How can we justify taking such a decision. A statement to the effect that there are no unknown relevant factors influencing the experimental situation is unfalsifiable, but it can be tested. Let us consider the case of the 'anomolous' notion of Mercury's perihelion, and the ceteris paribus clause used in this connection to isolate the Newtonian Gravitational Theory. Initially the testing of the ceteris paribus clause may be quite severe. From the theory itself we can calculate using the perturbations of Mercury's orbit, the path of a massive body which could account for the discrepency. This, infact, led in the nineteenth century to the prediction of the inferior mercurian planet 'Vulcan'. The ceteris paribus clause could thus be tested by observation. When Vulcan was not found, it was suggested that it was so close to the sun that it was, in practise, not detectable. But eclipse observations also failed to discover it.

As the 'obvious' explanations begin to run out, the testing of the ceteris paribus clause becomes less severe. The properties ascribed to the entity which causes the perturbation become more ad hoc. We could suggest, for example, that Vulcan was composed of very dense but transparent non-reflecting matter, (which accounts for its invisibility at night, during eclipses, and also during transits of the Sun's disc). An explanation in these terms is not unsatisfactory because it deviates from the framework of established astronomical theory which postulates that planetary material is all roughly similar, such predictions in the face of problems being of the utmost value. It is unsatisfactory because it does not stem from any unifying hypothesis and is not independently testable.

In general the severe tests of a ceteris paribus clause are those which involve the assumption of entities which form part of the theory's normal subject matter. The assumption of the presence of Vulcan did just this, as planets and masses are part of the subject matter of gravitational theory. There is, however, a limit to the type of assumption we can make without at the same time implying that the theory is inadequate to cope with the anomaly within the established theoretical framework. This is reached when the assumptions are of such a nature as to involve entities or processes which are outside the normal framework of the theory. But this does not imply that we assume at this stage a conjecture that the theory is false. It could be that the observed event was the result of two (or more) interactions, only one of which was explained (but adequately) by the theory. We might at this stage, for example, suggest that Mercury had a very strong magnetic field, and the perturbations were due to its proximity to the Sun's magnetic field. Again, however, we can test these assumptions more or less severely; more severely when our assumptions have independent implications, less severely when it is not the case that they have independent implications. When the imagination of the scientist becomes exhausted and all tests become merely repetitions, the severity of the tests declines drastically.

Although we can never be certain that our decision to accept the _ceteris paribus_ is correct, we can estimate by assessing the severity of the tests which have been proposed but not yet conducted when the decision is to be made. Frequently, when we can construct a model of the theory, we can use the model to _restrict_ possible _ceteris paribus_ assumptions. If we had a theory of enzyme action from which we could deduce (from the kinetics of a particular reaction) the charge on a certain part of the enzyme, but found that measurements indicated this charge to have a different value to the one predicted, a model of the mechanism of the reaction could restrict possible <u>ceteris paribus</u> assumptions. We might, for example, claim that the difference in the charges was due to the close proximity of a neutralising charge - but find on inspection of the model that this was ruled out because the spatial relations of the components of the molecule did not leave room for the presence of another charged atom.

Here we assume that the model is an accurate iconic representation of the molecule. If this assumption is correct, and the model shows that, in the region we are considering, there is no possibility of influence from other atoms, then a whole class of possibile assumptions which depend upon the presence of affecting atoms is ruled out. Again we get clues, but no proof, that there are no extraneous factors which affect the predicted charge. As for the relagation of observational theories to unproblematic background knowledge, the decision to accept the <u>ceteris paribus</u> clause can never be proved correct. We can submit our conjecture that it is to severe tests.

This discussion shows that, by careful testing, we can to some extent reduce the conventional element in making the decisions needed to falsify$_p$ a theory. We can never eliminate completely the possibility of error, but it can be reduced.

We have already given some reasons why falsification is essential for the maintenance of the empirical basis of a progression of theories, and can now clear up the question whether we need to falsify the theory system or the explanatory theory only of each stage. The notion of getting nearer to the truth has, as its logical counterpart the notion of advance from falsehood, but as we have seen the latter has methodological priority. To fix the point from which we can assess this advance we need an instance of falsification. If we are going to appraise explanatory theories in terms of verisimilitude we then need the instance to be falsification of a theory rather than a theory system. The logical content, and hence the difference in degree of verisimilitude of two theories is independent of any observational theory even though our fallible assessment of it may be so dependent.

If we are to estimate the difference in degree of of verisimilitude between two explanatory theories which require different observational theories to interpret their deductive consequences, we must be able to isolate the explanatory theory from the remainder of the theory system, otherwise we cannot conjecture that our assessments are of the truth or falsity content of the theory. Thus we need the conventional decisions needed to falsify$_P$ a theory in order for this to be possible. Whether we have correctly isolated the theory for the ascription of truth or falsity content, we shall never know, but by empirical experiment we can put it to the test.

(iii) Total Corroboration

The interpretation given so far to the appraisals do not give any indication that the preferred theories may be considered more reliable, in the sense that the future performances of the theory will be better than its predecessors. It may be argued that it is not the task of the pure scientist to arrive at reliable theories, and in one sense this is true. If he succeeded in achieving this goal to the extent that none of the theories in the body of scientific knowledge met with a counter instance for (say) twenty years, although we may agree that these theories are reliable, the pure scientist has failed in his task of modifying the theoretical framework in such a way as to restrict the number of logically possible states of affairs which are realised. If, however, we consider science to be 'a guide to life' we would consider the pure scientist's result as a great achievement.

Thus the pure scientist and the technologist are motivated by different goals, but nevertheless, the technologist expects that those theories which the pure scientist appraises highly will <u>also</u>, at least in many cases, serve his purposes better than others. It may happen that the technologist can find applications for false theories, but even in these cases, we can argue that calculations based on these theories approximate within acceptable limits, the results of the more highly appraised theories of the pure scientist.

We can give a conjectural estimate of the reliability of the pure scientist's theory by interpreting increasing corroboration as an increase in the degree of verisimilitude. If we have a series of theories each of which was <u>comparable</u> to the last, then we can assess the <u>total corroboration</u> of the final theory in the sequence as the sum of the excess corroboration of each theory over its predecessor in the series. We assume that each theory inherits the excess corroboration and increase in verisimilitude of its predecessor so that the final theory has a higher truth-content and a lower falsity-content than any other. If we interpret 'having the higher degree of verisimilitude' as 'being more reliable than' we can establish a partial ordering between theories with respect to their 'reliability'. That this ordering is only partial is obvious, as we cannot compare the excess corroboration between two theories unless one theory is comparable to the other, i.e. it represents a growth in empirical content with respect to the other theory. We cannot establish the ordering of two theories with respect to reliability unless they are members of the same sequence of theories. As Lakatos notes:

> "This circumstance reduces drastically the practical technological use of corroboration as an estimate of reliability for competing technological designs. For each design may be based on some theory which, in its own field, is the most advanced; therefore each such theory belongs, in its own right, to the 'body of technologically recommendable theories' . . ."[20]

Even such a limited notion of reliability is further handicapped by its fallibility. We can only conjecture that a more highly corroborated theory has higher verisimilitude, and, accordingly, our estimates of reliability will be equally fallible. Nevertheless, if we propose such a concept to give some justification for taking the most recent theories of science to be a better 'guide to life' than their predecessors we must necessarily use appraisals of excess corroboration which rely on type 3 comparisons.

To sum up, we have shown that by admitting the fallibility of the speculative assumptions used to interpret the tautological appraisals we can avoid the criticisms of Chapter Two when we compare theories with respect to their verisimilitude and their reliability. Such appraisals <u>are</u> fallible and restricted in scope. If we compare theories with respect to their verisimilitude we necessarily need a type 3 comparison to assess if one theory is more highly corroborated.

For methodological reasons, we also need a type 2 comparison by which we can state that the eliminated explanatory theory is false. Although we can give a theory of method without recourse to comparisons of reliability, one method of explaining why science is a 'guide to life' can be given in terms of an interpretation of 'total corroboration'. This again necessarily requires a type 3 comparison. Type 1 comparisons are needed to assess the axcess empirical content and degree of testability of a theory although the claim that absolute degrees of content can be measured by these comparsions seems dubious. Type 1 comparisons are also needed to assess the degree of corroboration of a theory, and have a related role in determining the unity of a particular series of theories. This we can show by considering two competing theories, T_1 and T_2, where T_2 is comparable to T_1. As was mentioned earlier, if we are to eliminate T_1 after some competition has lead us to preferentially choosing T_2 this asymmetric relation must hold. But this condition is only necessary, not sufficient. We also require that the excess explanatory content of T_2 is connected to T_1 in a way that is more than mere conjunction. If we have a theory T_1, 'All swans are white' which is refuted by the basic statement 'There is a black swan in Melbourne Zoo', then we would not eliminate T_1 on the proposed theory (T_2) 'All mute swans are white and all water boils at 100°C'.

Popper restricts such <u>ad hoc</u> theories by stipulating that the new facts should be <u>unlikely</u> or preferably of a <u>new kind</u>, i.e. inlikely in the light of all our total scientific knowledge. But this requirement is in some respects too strong and in others too weak. It is too weak becausedit does not rule out the possibility of adding a clause which, although unlikely in respect of all our scientific knowledge and although it leads to <u>successful</u> predictions, has nothing whatever to do with the content of the theory T_1. It is too strong because in many cases it is quite acceptable for the clause to predict facts which, although unlikely in the light of the previous theory T_1 may be predictions of highly corroborated theories in other fields. In this latter case we can sanction this manoevre if the conjunction of the (modified) T_1 plus the new clause represents a unification or synthesis of two previously unconnected fields of science.

Any criterion of what represents a satisfactory connection of two previously separated fields must of necessity be vague, but I would suggest that in some cases at least we can get some clues by an examination of the empirical content of the modified theory (i.e. T_1 altered in such a way as to remove the inconsistency with the accepted falsifier) and of the additional clause added to this. In one special case where the class of potential falsifiers of the modified theory and the additional clause had, as a subclass, the class of potential falsifiers of a third theory, we could immediately interpret this as an indication of a satisfactory unification, giving grounds for the elimination of T_1 and possibly also the third theory.

This is, however, a special case. More generally we would have to rely on a more intuitive criterion of satisfactory unification. Take for example, the refutation of the geological theory of mountain evolution: 'All mountains are formed by volcanic activity' by the accepted statement 'There are many mountains which are formed of sedimentary rock' (hence not formed by volcanic action). This can be satisfactorily replaced by the new theory : 'All "hot" mountains are formed by volcanic activity and all "cold" mountains are formed by the collision of two tectonic plates'. Although this added clause is independent of the volcanic theory and predicts facts which are unlikely in the light of it, these facts are *not* unlikely in the light of previous geological theories of the expanding earth.

If we call the volcanic theory 'T' and the modified theory plus the added clause 'T^*', then the potential falsifiers of T would include such statements as 'X is a mountain which is not volcanic'; the potential falsifiers of T^* would include such statements as 'Y is a "cold" mountain which is not formed by tectonic plate movement'. Thus, although the falsifiers of T are not a subclass of the falsifiers of T^*_* we could, by an *a priori* comparison establish a correspondence between them - if we assume that all mountains are either "hot" or "cold". Then for each potential falsifier of

T there would correspond two potential falsifiers of T^*. If, on some intuitively acceptable assumption, we can draw up a correspondence between the falsifiers in such a way, we can assess the satisfactoriness of the unification.

This sort of comparison is by no means generally applicable and frequently we must rely on a much more speculative assumption in claiming that the unification is justified. Often _a priori_ comparisons are useless; we may rely on the empirical success of the new theory itself to unify in our conceptual framework previously disparate phenomena.

Now we are in a position to give the full requirements for the falsification and elimination of a theory from the body of scientific theories. In doing so, I shall clarify one or two points in the slightly over-simplified account of the NMF theory of method given in the introduction.

An (explanatory) theory is rejected as 'false' and untrustworthy if a potential falsifier of the theory has been accepted in the light of observational theories conjectured to be reliable. If a _ceteris paribus_ clause is required for falsification, this must have been severely tested. An explanatory theory T_1 is eliminated if a theory T_2 has been proposed such that:

(1) T_2 is comparable to T_1.

(2) T_2 has excess independent content over T_1

(3) This content is more intimately related to the content of T_1 than by mere conjunction.

(4) Part of this content has been 'verified' by experiment.

It follows from these requirements that (a) if a theory has been falsified it will not necessarily be eliminated, and (b) if a theory has been eliminated it is not necessarily false. BUT coupled with these requirements are two methodological 'directives'.

(1) If a theory has been falsified, all possible haste must be made to replace it by one which eill enable the requirements for elimination tobe met. In most cases, as Popper points out, this directive will be gratuitous, for, "we will have, before falsifying a hypothesis, another one up our sleeves; for the falsigying experiment is usually a crucial experiment designed to decide between the two. That is to say, it is suggested by the fact that the two hypotheses differ in some respect, and it makes use of this difference to refute one of them". It is, however, possible to falsify a theory without such a replacement, in which case the body of our scientific knowledge will be inconsistent as it will contain a deductive consequence of the falsified theory and the accepted basic statement.

(2) The second directive is that we should make all reasonable attempts to falsify T_1 if T_2 is accepted as a superceeding theory, and the more frequently this occurs in any series of theories, the harder we should try. In many cases this directive also will be gratuitous; the experiments which corroborate the new predictions will frequently lead to <u>questioning</u> of the capabilities of the old theory, and frequently will lead directly to the acceptance of basic statements which falsify it.

I should stress that this interpretation of NMF differs in some respects from Popper's, although I believe it is true to the spirit of his doctrine. The overriding principle which characterises the whole of this theory of method is that criticism of, and hence competition between theories should take place as frequently and as vigourously as possible.

CHAPTER FOUR

In this chapter I shall criticise the 'sophisticated methodological falsificationism' (SMF) theory of method proposed by Lakatos. I shall do this because this theory imposes certain limitations on the use of comparisons which I claim to be essential. It lacks completely both type 2 comparisons, and those type 1 comparisons which assign _a priori_ measure functions; but also, because of the demarcation principle used, the status of _any_ _a priori_ comparison between _theories_ or theory systems, become highly uncertain.

As I claim that all these comparisons are essential to any satisfactory account of scientific progress, I must demonstratethat SMF is unsatisfactory. Moreover I believe that the shortcomings of SMF are representitive of any theory of method which restricts comparison in this way. It would be difficult to argue this case conclusively, but we can put forward several difficulties which would face any theory of method which relies, as SMF does, exclusively on comparisons of progressiveness, (for the _a priori_ element q.v. p.122).

My criticisms will be of two kinds. In section A, I shall call attention to several _practical_difficulties which would face the scientist if hw accepted SMF as a prescriptive theory of method. In Section B, I shallw try to show that the apparent methodological advantages of SMF over NMF (of Chapter Three) <u>are</u> only apparent, and that the adoption of the SMF demarcation principle and falsification$_L$ requirements (1) does not solve the problems facing NMF, and (2) leads to new problems peculiar to SMF.

Firstly, let me remove any confusion as to which doctrine I refer to by 'NMF'. Ladatos himself directs criticism towards a theory of method which he calls 'naive methodological falsificationism' but it is only a highly modified version of <u>this</u> theory which I wish to call 'NMF' and uphold against SMF. (Lakatos' 'NMF' is characterised by the methodological directive that when a theory has been falsified$_P$ it must be eliminated immediately. The body of scientific knowledge in current use is not allowed to become inconsistent, and hence, presumably, could contain 'voids' where a falsified (eliminated) theory had not been replaced by one which was comparable to it. This type of NMF also lacks methodological rules which restrict the '<u>ad hoc</u>' introduction of new theories.)

The theory of method which I shall call 'NMF' is that given in Chapter Three. It is, in many respects, similar to SMF. Both are non-justificationist theories of method which rely heavily on the notion that the empirical character of scientific knowledge depends upon the method of its growth. Taken in isolation from the 'methodological directive', the elimination requirements of NMF are identical with the falsification$_L$ requirements of SMF - at least in the case where a 'progressive theory-shift' (q.v.) results from the replacement of an explanatory theory. As, in NMF, we <u>can</u> eliminate a theory without having falsified$_P$ it, it may be objected that I am just splitting hairs in trying to distinguish between the two. This, however, is not the case. The two theories of method stem from radically different temperaments which I hope to make clear during this chapter.

The transition from NMF to SMF involves a semantic re-interpretation of 'falsification'. In NMF, falsification$_P$ involves an experimental result which we decide conflicts with the theory under test. Falsification$_P$ amounts to the acceptance of corroborated counter-evidence, a relation between a theory and the empirical basis. Thus, in the theory system under test, the relevant conventional decisions isolate the <u>explanatory theory</u> and it is declared in conflict with our 'unproblematic' background knowledge.

Although this refutation does not lead to its elimination from the body of scientific knowledge it leads the way to a substitution into the theory system of a new explanatory which enables the elimination conditions for the the old theory to be met. (If the introduction of the new explanatory theory calls for a revision of the observational theories of the old system, these new OT's must be acceptable as unproblematic.) In NMF we admit that science <u>can</u> grow without refutations leading the way - by the elimination of an unfalsified$_P$ theory - but the methodological directives stipulate that this must not happen frequently.

In SMF we make no conventional decisions to isolate the explanatory theory from the rest of the theory system. The term 'counter evidence' "has to be abandoned in the sense that no experimental result must be interpreted as counter evidence" for a particular explanatory theory. In SMF we regard an explanatory theory as falsified if and only if it has been replaced in a thoery system such that the new theory system (1) has excess corroborated novel empirical content over the old, and (2) the unrefuted content of the old system is explained in terms of the new.

However, 'counter evidence' for an old theory system is not directed towards the explanatory theory. In SMF "it is not that we propose a theory and Nature may shout NO; rather we propose a maze of theories (a theory system) and Nature may shout INCONSISTENT".

To repair this inconsistency, we do not immediately try to replace the explanatory theory. "We never reject a specific theory simply by fiat. If we have an inconsistency ... we do not have to regardwhich ingredients of the theory (system) we regard as problematic and which as unproblematic: we regard all ingredients as problematic in the light of the conflicting, accepted basic statement, nad try to replace all of them". Which parts of this mutually inconsistent theory system should be replaced ? "The sophisticated falsificationist can answer that question easily ... Try to replace first one, and then the other, then possibly both, and opt fot the new set up which provides the most progressive problem shift". (A 'progressive problem shift' is a move from onw theory system TS_1 to another TS_2 such that TS_2 predicts new facts; a problem shift is empirically progressive if the new facts are corroborated).

Gone in SMF is the methodological directives to attempt falsification$_P$ of any theory which has been eliminated. Gone also is the directive that all haste must be made to replace a falsified$_P$ explanatory theory.

SECTION A

A criticism based on practical difficulties only holds water if the proposed methodology claims to be prescriptive. Even though both SMF and NMF contain descriptive elements, I shall assume that their primary aim is characterising a method which it is rational for the scientist to adopt. Without any theory of method the procedure of the scientist will be changed by trual and error if he finds that one approach enables him to achieve his long-term goals more readily. As he is thus likely to end up with a 'practical' methodology which is successful, it is not surprising that many elements in this practical methodology will be paralleled by similar compnents in a rational theory. But when we decide between competing theories of method, I shall assume that of an accurate description of the methods of the scientist will only be a reason for choosing a theory if it can be shown that these elements have a rational justification within the framework of the theory. If, within this restriction, a particular theory of method describes more accurately the actual practise of scientists than another than I shall look upon this merely as an added bonus.

This is not to say that we should frame theories of method in isolation from a study of the practise of science; obviously many of the practical methods developed will be capable of rational reconstruction simply because of the close connection between rationality and success. But also we look to the scientific enterprise itself to determine what a prescriptive methodology must achieve if it is to be considered successful. (A successful methodology must be one which, if adhered to, will enable the enterprise to realise its own goals.)

I must state here that my first criticism of SMF is based on am assumption connected with the goal of science, Of course, we can only speak metaphorically in ascribing goals to the scientific enterprise itself, and we could argue that these goals will be as many and varied as there are individual scientists.

> "And yet" as Popper states, "we do feel, more or less clearly, that there is something characteristic of scientific activity; and since scientific activity looks pretty much like a rational activity and since a rational activity must have some aim, the attempt to describe the aim of science may not be entirely hopeless."[21]

I think it may be asserted that the aim of science is to arrive at theories of ever increasing explanatory power, and which, we conjecture, are nearer to the truth than their predecessors. This aim, I will call the aim of <u>growth</u>. My assumption concerns the <u>rate</u> of growth. I believe that adherence to a methodology which, in a shorter period of time produces the same, or greater growth, is more rational than adherence to one which is slower. In other words the aim of science is not only growth, but <u>rapid</u> growth.

Lakatos claims. for SMF, that because it involves fewer fallible conventional decisions (those involved in 'isolating' the explanatory theory from the background knowledge), this means there will be fewer opportunities for <u>error</u>, and hence os the <u>rational choice</u> is for that methodology which will less probably lead us away from the truth we should adopt SMF rather than NMF as a prescriptive theory. This claim although partly true is misleading. We <u>do</u> choose between theories of method by assessing which is the less likely to lead us into error. However, in another sense, the least fallible theory might not necessarily be the most rational to adopt. It may well be (with one important <u>proviso</u>) that a more fallible methodology which prescribes a bolder, perhaps more speculative, but <u>rapidly accomplished</u> programme, will enable us to achieve more rapid growth than a more cautious methodology which, at each step in the programme may be less likely to lead us into error.

The <u>proviso</u> is that, following the rapidly accomplished programme, we must always be prepared to admit our errors (and this presupposes some method of detecting them) and learn from them. If the faster programme takes three steps forward and one step back for each single step taken by the slower programme, then, if the goal is rapid growth it is more rational to choose the former. This argument presupposes that the 'quality' of the growth is the same in each case, but as both NMF and SMF rely on a similar 'quality control', that of our novel predictions being corroborated, this condition will be satisfied.

What then are the practical diffieculties associated with the methodological directive of SMF when a scientist is faced with an inconsistent theory system? He can no longer regard an experimental counter example as refuting the theory he is testing. He must look at <u>all</u> parts of the theory system involved in the test and see if replacement of <u>any</u> of these components will lead to a progressive probelm shift. Especially suspect will be the observational theories in the light of which the decision to accept the 'potential falsifier of the theory' was taken. If this observational theory could be given a semantic interpretation or have added an auxiliary clause which made a new theory system (consisting of the old explanatory theory plus a new observational theory) once more consistent (but also progressive) then the old observational theory would be declared false$_L$.

The first difficulty I wish to point out is the problem of <u>specialisation</u>. The limitations of the human brain and the inclinations of a particular scientist are such that in the increasingly complex enterprise of science a group of scientists who are involved in a particular research programme may have little or no idea of the subtleties and intricacies of the obwervational theories which they emply to interpret their data. This is particularly true of fields of science which rely heavily on complex instrumentation. A geologist, for example, may analyse a sample in which he predicts the presence of certain trace elements using a mass spectrograph. If the machine produces results which are not consistent with his predictions, then he <u>may</u> assume that the machine is not working properly, but it seems rather a lot to ask of him to bring into question the observational theory by which he interpreted the data as indicating the presence of certain elements which he predicted were absent. This observational theory - a theory of atomic physics predicting the trajectories of ionized particles moving in strong magnetic fields - is hardly a subject in which the geologist is 'well up'. Although this may be somthing of an extreme example it is by no means an uncommon phenomenon.

Lakatos' account of SMF seems to imply a 'division of labour' between 'theoreticians' and the 'experimentalists'. The experimentalist comes up with a result which falsifies$_P$ an explanatory. Then the "theoretician appeals against the verdict of the experimentalist ... (questioning) the interpretitive theory in the light of which its truth-value had been established". To what extent such a division of labour takes place, I suppose is a matter for some sociological scientific survey - but even if we had some such procedure whereby the mambers of the explanatory research programme could go th their 'theoretical' collegues to inquire about the possibility of alterations in the observational theory to (1) resolve the inconsistency and (2) generate a progressive problem shift, we are <u>still</u> faced with a second practical problem.

This second difficulty is the problem of the asynchronous growth of theories. By this I mean the problem raised by the fact that the periods of theoretical advance where the time is ripe for a change in theoretical outlook, where, perhaps, several rival theories are in competition, do not come to all theories of a theory system at the same time.

If now an experimentalist produces a 'counter example' to an explanatory theory when the momentum of development of the observational theories was at a low ebb, the theoretician may have to wait for a considerable time before he can be satisfied that no new observational theory is forthcoming to reconcile the inconsistency. But just how long does he have to wait before he takes the decision that a change in the explanatory theory is called for ? SMF provides no answer.

The retreat to a 'pluralistic model of the deductive structure of theories' where we consider the possibility of change in <u>any</u> part of the theory system is accompanied by the very real danger that the proponent of a particular explanatory theory can prevaricate and procrastinate the search for a new explanatory theory. If we do not stipulate some constraining time limit which determines when an explanatory theory must be rejected, then SMF could degenerate into the philosophy of scientific method advocated by Freud.

> "None of us can guess what the ultimate judgements about our theoretical efforts of mankind will be. There are instances in which rejection by the first three generations has been corrected by the succeeding one and changed into recognition. After a man has listened carefully to the voice of criticism in himself and has paid some attention to the criticisms of his opponents, there is nothing for him to do but with all his strength to maintain his own convictions..." 22

However, behind this 'pluralistic' doctrine of SMF there lies a genuine problem. Sometimes experimental results which are inconsistent with an explanatory theory <u>are</u> overthrown by a change in observational theory leading to a re-interpretation of the data. Lakatos gives us an example; the Proutians fight against the Stasian refutation.

> "For decades Prout's theory T (that all atoms are compounds of hydrogen atoms and thus "atomic weights" of all chemical elements must be expressible as whole numbers) and falsifying 'observational' hypotheses, like Stas's 'refutation' R (the atomic weight of chlorine is 35.5) confronted each other. As we know, in the end T prevailed over R" 23

Lakatos shows that T and R can be considered inconsistent by putting them into the form: T = 'the atomic weight of all pure elements i.e. homogeneous elements, are multiples of the atomic weight of hydrogen'; and R = 'chlorine is a pure (homogeneous) element and its atomic weight is 35.5'. Hence a statement of the form 'At x,y,z,t (some spatio-temporal coordinates) there is a pure sample of chlorine with atomic weight 35.5' would be an accepted basic statement which, backed up by the falsifying hypothesis R, is a counter-example to the theory T. But, asks Lakatos, how well corroborated is R?

> "Let us have a closer look at the fine structure of 'chlorine is a pure (homogeneous) chemical element' (R_1). In fact R_1 stands for a conjunction of two longer statements T_1 and T_2. The first statement, T_1 could be this: 'If seventeen chemical purifying procedures p_1 p_2 . . . p_{17} are applied to a gas, what remains will be pure chlorine.' T_2 is then: 'X was subjected to the seventeen procedures p_1 p_2 . . . p_{17}.'"[24]

The experimentor now applies all the purifying procedures but the conclusion that what remains must be pure chlorine will only be a 'hard fact' in virtue of T_1. Lakatos calls T_1 'the interpretative theory' in the light of which the experimentor decides that the chlorine is pure. (In fact T_1 has more the form of a <u>definition</u> of purity, than a theory of purity, but let us accept that this stems from a physico-chemical theory which we will call T_1.)

The problem which faces us is that a change in T_1, involving the addition of p_{18}, an isotope-separation process, to the other purification procedures, will lead us to reassess the claim that what remains after the sevevteen purifying procedures <u>must</u> be pure chlorine. Thus although R as interpreted by T_1 is a falsifying hypothesis of T, R as interpreted by T_1' (i.e. T_1 plus procedure p_{18})is not. The 'facts' have been re-interpreted and the original inconsistency between the explanatory theory and the 'counter-evidence' removed by this change in the observational theory.

What is the NMF attitude to such a 'saving' of the explanatory theory by the introduction of the new observational theory? As I interpret this theory of method there are no problems at all. When the theory T is falsified by R in the light of the successful observational theory T_1, T is rejected as false. Immediately we search for a new explanatory theories which could lead to the elimination of T. We propose, for example, the theory "All atoms are composed of an integral number of particles which have half the mass of hydrogen atoms". We devise tests for these new theories - in this case it may involve attempts to split the hydrogen atom into its component parts. (Who knows what difference might have been made to the history of the world if NMF had been the prescriptive methodology in the nineteenth century !) With hindsight of course,we can say that these theories would not have been successful, but what we cannot say is that the testing of them would not have led to new and interesting problems being uncovered.

Now every observational theory is, in its own context, an explanatory theory. The above mentioned physico-chemical theory T_1 became, in the hands of Crookes[25], a theory which he used to deduce certain observation statements predicting the phosphorescent spectra of the rare-earth yttrium after it had been subjected to continued chemical purification. From this physico-chemical 'theory of purity' we can predict that any purification procedure which produces samples of yttrium which have the same characteristic spark spectra produce pure (homogeneous) samples of yttrium with the same phosphorescent spectra. This prediction was refuted by Crookes who demonstrated a fraction procedure which resulted in samples of yttrium having the same spark spectra (and hence the same element) but with differing phosphorescent spectra (hence indicating that the yttrium atoms were not identical). This refutation led eventually to the elimination of T_1 and its replacement by T_1' and the recognition of atomic isotopes. This theory system involved an observational theory by which a spark discharge spectra could be interpreted as characteristic of a particular atomic structure.

The SMF methodologicst could now, of course, challenge <u>this</u> observational theory, and question whether the 'counter-evidence' against the physico-chemical 'purity theory' could not be progressively resolved by a change in the observational theory of atomic spectra.

But the NMF methodologist, when the T_1' theory is corroborated by tests, accepts that he was <u>in error</u> in taking part of his background knowledge as unproblematic, and in claiming T to be false. Now he takes T_1' as unproblematic and, unless he has eliminated T in the meantime, once more accepts it for continued testing.

Thus both SMF and NMF have an 'appeal procedure' but, reflecting the different temperaments of the two theories, different emphasis is placed on the role of the scientist. In NMF the scientist takes any observational theory which is well corroborated as unproblematic <u>until</u> that theory in an explanatory role is refuted - probably in some different branch of science. In SMF the scientist questions the observational theories of the theory system he is using immediately to see if they can be progressively replaced. But even the SMF scientist can only <u>delay</u> the decision to accept an observational theory as unproblematic, otherwise he will never be able to make any decisions at all. Thus the claim that SMF does not need decisions to isolate unproblematic background is rather empty. All that is done is to <u>defer</u> the decision to accept an observational theory.

Let us look at the performance of the scientist who adopts NMF from the standards set by SMF. If we have the case where the SMF method decides after an examination of the inconsistent theory system that a replacement of the explanatory theory is progressive - the NMF scientist would have done that straight away. If we have the case where the SMF method decides that the replacement of the observational theory is progressive then the NMF scientist will adopt the new observational theory when corroborated. In the meantime, however, he would have been testing new explanatory theories which may have led to new problems.

On one condition the NMF scientist cannot loose - he can only gain. This condition is that flaws in 'observational' theories are detected by <u>other</u> scientists who use them in an explanatory capacity. This is not an unreasonable assumption to make. The NMF scientist, because he is prepared to make dogmatic (but in some cases retractable) assertions about the <u>facts</u> has fewer (but perhaps more fallible) decisions to make in deciding how to overcome a 'counter example'. If he does have to retract his assertion he will be at <u>worst</u> in the same position as the SMF scientist. If he does <u>not</u> he will have a temporal advantage over the SMF scientist, for he will have immediately rejected the explanatory theory and, if he does not have one already, he will be looking for a replacement, while the SMF scientist must see if a replacement of the explanatory theory, or of the observational theory or both will lead to a progressive problem shift.

Thus both NMF and SMF can cope with the problem behind the 'pluralistic' doctrine of SMF, but NMF if anything slightly better. We have, however, pointed out two practical difficulties in replacing the observational theories of a theory system, which would suggest that this theory of method would, in many cases, lead to a slow development of new thoery systems. A similar argument applies to replacements of the _ceteris paribus_ clause. In SMF the _ceteris paribus_ is also considered to be a potentially problematic component of the theory system - and once again we have to decide whether a replacement would lead to the creation of a progressive problem shift.

The practical difficulties of SMF are, then, (1) the large number of possible alterations to the theory system, each of which needs to be assessed in terms of progressiveness, leads to slow progress from theory system to theory system, and (2) in some cases, unfamiliarity with, or unavailability of, potential replacements leads to further delay.

SECTION B

In this section I shall consider the methodological implications of the SMF demarcation criterion and the requirements for falsification_L.

So far, my account of the demarcation principle of SMF has been slightly oversimplified. We have given the criterion by which we determine whether a particular theory system is scientific. But one characteristic feature of SMF is the stress itplaces on series of theories. The criterion of scientific acceptability for series of theory systems is rather different.

A series of theory systems TS_1, TS_2, TS_3 is one where each seccessive theory system (e.g. the addition of an auxiliary clause, or the semantic reinterpretation of one of the component theories) which produces an increase in the empirical content, (that is, prediction of some new fact). A series of theory systems is called 'theoretically progressive' if this condition is satisfied. If for each theory system part of this excess content is corroborated, then the series is called 'empirically progressive'. A series of theory systems is called 'progressive' if it is both theor-etically and empirically progressive, and degenerating if it is not.

The demarcation principle for a series of TS demands that for a series of theories to be 'scientific' it must be at least theoretically progressive. In other words, SMF does not demand that <u>each</u> superceeding theory system must have had its excess content corroborated, for the series to be considered 'scientific'. (It is not just a question of the scientist not carrying out the experiments which might have led to corroboration - as long as the TS is theoretically progressive it may be <u>refuted</u> by experiment.)

The NMF requirement for acceptability of a new theory is not only that it has increased content, but also that this excess content <u>must</u> be corroborated, and we have a similar requirement for the SMF demarcation principle for <u>particular</u> theory systems. However, when we consider a <u>series</u> of SMF theory systems, it is permitted to consider such a series 'scientific' even if it contains theory systems which do not satisfy the requirement of corroborated content. This is a concession on the part of SMF to the phenomenon of the '<u>tenacity of theories</u>'. In particular research programmes, a scientist will continue to use an inconsistent theory system hoping that by various modifications he will, eventually be able to achieve a theory which is well-corroborated.

This concession to the tenacity of a theory system in the face of refutation has parallels to the NMF concept of the tenacity of an empirical theory in the face of falsification$_P$ (i.e. falsification is not a ground for elimination). Moreover, in SMF, there is no methodological directive to attempt to replace the explanatory theory, as there is in NMF. Rather we accept that we must replace the theory system by another, but even now, this superceeding TS, although it must have increasing content, need not necessarily have excess <u>corroborated</u> content.

There is, however, in SMF, a methodological directive: "Intermittantly the increase in content must be retrospectively corroborated". When this happens the theories which are superceeded by the theory system with the higher corroborated content are falsified$_L$ and eliminated. This contrasts radically with the NMF requirement that a new explanatory theory <u>must pass</u> the <u>first</u> test of excess content if it is not to be rejected. In SMF the explanatory theory of a TS which failed the first corroborating test is not necessarily rejected ("It was careless of Popper to attach so much importance to the first test") so long as we can produce a modification of the TS so as to (1) remove the inconsistency, and (2) produce a theoretically progressive new theory system which includes the old explanatory theory. Even this new theory system need not be experimentally corroborated for the series to count as scientific.

The extremely unsatisfactory feature of this demarcation of 'scientific' series of theories, is that we have no means of knowing <u>when</u> a long period of 'theor-etical progress' will result finally in a TS which is corroborated. If no time limit at <u>all</u> is given after which the whole series is to be abandoned, then we face the poosibility of merely a perpetual series of <u>ad hoc</u> manoeevres to restore consistency of the previous theory system in the series. For at any stage in the series we can claim 'It will be the <u>next</u> theory system which is empirically corroborated'. Thus, although it <u>may</u> have been careless of Popper to attach so much importance to the first test, it seems that we must a attach importance to at least some time limiting factor to prevent the type of degeneration outlined above. The SMF requirement of 'intermittant corroboration' seems too weak to prevent such <u>ad hoc</u> manoevres. The bolder approach of NMF may lead to the rejection of a 'true' theory when it fails a corroborating test due to a false observational theory, but, as we mentioned before, this can be rectified when a new observational theory is corroborated. In SMF we are less likely to throw out a good theory but this is at the cost of prolonged and possibly fruitless attempts at modification of the theory systems to produce eventual corroboration.

In NMF of course, we require that each new theory which is put forward has part of its novel empirical content corroborated at the first test, and it is because of this, and the fact that the superceeded theory is frequently falsified, that we can claim to interpret this type of empirical growth as increasing verisimilitude. When we compare these stringent requirements to those of SMF we see that a secondary interpretation of SMF appraisals to explain why succeeding theory systems can be considered nearer the truth, is on very much weaker ground.

Not only do the SMF appraisals completely lack the notion of falsification$_P$ of the superceeding theory (this is paralleled only by 'an anomoly of the previous theory system') but also the notion of interpreting 'successful predictions of new facts' as 'indication of an increase in the truth content of the theory system' is applicable, not as a matter of course, but only at 'intermittant intervals'. Now it may well be, to take a simple case, that a series of theory systems involves changes in both explanatory and observational theories. Let 'ET' stand for 'explanatory theory' and 'OT' for 'observational theory'. We will mark with * those theory systems which have been corroborated. Consider the series:

$TS_1(ET_1 + OT_1)^*$, $TS_2(ET_2 + OT_2)$, $TS_3(ET_3 + OT_2)$, $TS_4(ET_3 + OT_3)^*$

then, on the basis of the excess corroborated content of TS_4 over TS_1 we may agree to interpret this in terms of a conjectural assessment of increasing truth content of TS_4 over TS_1. But it is difficult to see what grounds we have for claiming that TS_3 has anyincreased truth content over TS_2. Both TS_2 and TS_3 are inconsistent and have no corroborated excess content. It seems that in this case, SMF, permits the replacement of one theory for another in a situation where we have no grounds for even conjecturing that the superceeding theory is nearer the truth.

We argued on p 81, that to preserve a firm empirical basis for the claim of increasing verisimilitude we require that falsification$_P$ was essential in order that we may learn from our mistakes. It may be objected that the corroboration of the excess content of a theory, (which falsifies$_L$ the preceeding theories) is sufficient to establish retrospectively the empirical basis fo the whole series of theory systems which have preceeded it. Each TS is proposed to overcome an inconsistency in the previous system and this is sufficent to isolate the problem needed to establish the empirical foundation for progress. We may reply to this (1) that it is not necessary for any anomaly or inconsistency to be present in a theory system for it to be falsified$_L$. All that is required is that a new theory system be proposed with a higher corroborated content. (2) Even if falsification$_L$ was always of a theory system which had been shown to be inconsistent, as we cannot direct the arrow of the modus tollens to any particular component of the theory system, we have no method of determining whether a new theory system which resolves the old inconsistency has dealt with the origins of the problems which faced us.

It may, of course, be argued that a resolution of inconsistency is all that we can expect to achieve - but there is a danger that this view will lead to instrumentalism.

Another shortcoming of SMF which results from the pluralistic doctrine, is that we cannot conduct comparisons between explanatory theories which rely upon different interpretitive observational theories. As mentioned on p 116, if we are to be able to conjecture that our assessments of truth and falsity-content are of an explanatory theory, we must be able to isolate that theory from the remainder of the theory system, which for this assessment we take to be unproblematic.

In SMF, however, any estimate of the difference of degree of verisimilitude can only be assessed between two theory systems unless we take the conventional decision to consider the observational theories unproblematic - a move quite foreign to the spirit of the pluralistic theory of deductive structure.

Thus we see that the restriction of the use of conventional decisions which enable us to interpret our observations as 'unproblematic' facts, and hence prevent us from undertaking comparisons with 'the facts' leads to severe difficulties; in (1) a proliferation of the number of possible ways of restoring the consistency of a theory system, (2) a weaker justification of the interpretation of appraisals in terms of verisimilitude, (3) a loss of the empirical basis provided by falsification$_P$. But also the demarcation principle for series of theories does not satisfactorily guard against continued '<u>ad hoc</u>'non-corroborated growth.

Let us examine more closely the demarcation principle for theoy systems. It is related to the demarcation of 'scientific' series of theories by the requirement that for a theory system to be 'scientific' it must be a corroborated member of a series of theories. A rather strange result which follows from this is that all theory systems of a 'scientific' series of theory systems need not be 'scientific'. But there are more serious objections to the intrinsic nature of this demarcation principle, depending as it does upon a retrospective classification of a statement as scientific, only when it has 'proved its mettle'. We may briefly recontruct a hypothetical 'evolution' of the SMF demarcation principle to show these objections.

The positivists' demarcation criterion was:

(1) "All statements which are non-analytic, and which are true or false are meaningful and scientific", (the rest are 'metaphysical$_1$' i.e. nonsense.)

Popper's demarcation principle was:

(2) "All statements which can be interpreted as falsifiable are scientific", (the rest are 'metaphysical$_2$' i.e. non-science + nonsense.)

Lakatos' demarcation principle (for theory systems) was:

(3) "All statements which have excess corroborated empirical content over their predecessors are scientific", (the rest are 'metaphysical$_3$' i.e. poor science + non-science + nonsense.)

(Lakatos, of course, would object to the use of 'metaphysical$_3$' in this way.)

The point I am trying to make is that the successive demarcation criteria are restricting the range of criticism which can be applied to 'scientific' statements by more restrictive conditions which a statement must satisfy in order to be 'scientific'. We cannot begin to criticise scientific statements qua theories, until they pass the standard set by the demarcation criteria. For example, Lakatos' demarcation principle rules out a priori comparisons of falsifiable statements until they have been tested and shown to have corroborated excess content. As a warning to those who wish to carry this procedure further, let me propose 'Perry's demarcation principle':

(4) "All statements are scientific which have excess corroborated empirical content over their predecessors, and which cause such criticism to be directed against them that they are eliminated and replaced with a statement which has even more excess corroborated empirical content, and which causes such criticism . . ."

In other words, I believe that, the most satisfactory demarcation principles are those which most quickly isolate just what it is that is the subject of our criticisms.

L

Lakatos' demarcation principle fails to do this, as we can only determine what is and what is not a theory _after_ we have conducted empirical tests. But why, then, should we not choose the positivists' demarcation principle - for this would isolate 'scientific' statements most rapidly of all? The answer is that, depending as it does on the notoriously difficult problem of meaningfulness the criterion is not sufficiently precise and it is difficult to get any intersubjective agreement as to a method of assessing 'sense'. Why do we choose Popper's demarcation principle over Lakatos'? In SMF, the demarcation of a _statement_ as scientific corresponds to the empirical testing of a _theory_ in NMF. But in NMF this theory was already, in virtue of its falsifiability, a scientific statement, which could be _criticised_ and compared with other theories by _a priori_ comparisons. The SMF demarcation restricts the comparisons we can make between, and criticisms we can make of statements _qua_ scientific statements. Because the principle makes the notion of a 'single scientific theory' a catagorical mistake (unless we are speaking of _a_ member of a series of theories) we cannot make theoretical appraisals by any comparison which assign absolute measure functions such as degree of simplicity or degree of probability, or indeed ascription of truth or falsity.

In attempting to 'pre-select' only the 'good' theories from among the class of statements, the SMF methodology is reminiscent of the biblical parable of the wheat and the tares. On finding their masters' field infested with tares, the servants asked if they should go and pick them out. The master told them to wait until harvest time; they could separate them then. In a similar way the NMF methodology accepts all theories, good and bad, and decides between them by criticism based on all three methods of comparison. The SMF methodology 'pre-selects' only those theories which NMF would call corroborated. I am suggesting that this effectively limits the type of criticism which we can put forward and hence limits also the rate of growth of our knowledge.

CONCLUSION

In this essay we have examined some aspects of the problem of the growth of knowledge from the point of view of comparisons which can take place between competing theories. The preferences which resulted from comparisons relevant to scientific theories were found to be objective preferences. The relation 'S prefers X to Y' as applied to theory preference can only vacuously be interpreted as 'A person S prefers X to Y'; S in these cases represents the 'impersonal scientific community', which lays down goal directed objective standards of betterness.

We introduced three methods of comparison and claimed that any adequate account of scientific progress must contain representitives of each of these methods. The interest of such a claim depends on just what we take to be the requirements of adequacy for this adequate account. I suggest these include:

(1) Some explanation as to how we arrive at theories, and how we decide from the infinitely many possible theories which ones we choose to test.

(2) Some account of how we can rate the performance of the theories we regard to be test-worthy with respect to their capabilities of adding to our knowledge. Also

we need some account of how we can determine when a new theory has satisfactorily replaced theold in its functions.
(3) We need some account as to why it is rational at least to conjecture that our superceeding theories are nearer to the truth and mor reliable than their predecessors.
(1) To satisfy the first requirement it is obvious that a priori comparisons must by employed. As we have no information concerning the empirical testing of these theories any method of selection between theories prior to testing must be a priori.

Although this is a question quite foreign to the falsificationist methofologies which I have been primarily discussing, it seems to me to be quite legitimate for a philosopher of science to ask for some method of rationally reconstructing the scientist's arrival at a particular theory. This question is usually ruled 'out of court' by the falsificationist, or fobbed off by the following typical arguement. "The creation of a scientific theory is just that, a creation of the mind - in the same way as a poem or the score of a symphony. We do not ask 'How did Beethoven arrive at the Sixth Symphony ?' and even if we did the answer to these questions must be purely psychological and not the province of logico-philosophical enquiry". But this analogy is not well founded; although there certainly is a creative element involved the arrival of a scientist at a particular theory, he does not have the implied unlimited freedom to come up with just whatever theory takes his fancy. Essentially he is bound by the problem which he is attempting to solve. If we had to reconstruct how a scientist arrives at his theories, I cna see no reason, in many cases, why we should not accept Peirce's notion of 'abduction'.

> "All the ideas of science come to it by the way of Abduction (consisting) in studying the facts and devising a theory to explain them. Its only justification is that if we are ever to understand things at all, it must be in that way ....Abduction merely suggests that something may be." 26

All our postulate hypotheses are just 'risky guesses' but a scientist does not try out every wild idea that enters his head. Of the vast majority of them, he will say when asked 'Why did you not try that one?' 'It did not seem reasonable - it provided no explanation for the problem I was considering'. In other words we can reconstruct how many of these guesses could be eliminated by means of an _a priori_ comparison.

(2) We gave accounts of how we could rate the performances of theories with respect to testing; the probabilistic account by which, if we allowed an assessment of the individual predictions of the theory to count as an

appraisal of the theory, we could rate theories in the light of confirming evidence. But we raised some difficulties with this approach. Nevertheless, if we are going to claim that a superceeding theory has taken over the functions of an old theory in so far as it describes or explains at least all the facts which are explained by the old theory or in so fat as it is confirmed by the evidence at least to the same extent, then we will necessarily require type 2 comparisons to determine wherher this is so.

But type 1 and type 2 comparisons by themselves can offer no solution as to how to critically restrict the new theories which we propose to be free from content decreasing stratagems. If we require to prevent this the new theory must in some way be shown to have excess empirical content over its predecessors, Type 2 comparisons cannot appraise excess content; type 1 comparisons can, but they cannot appraise whether the excess content is empirically realised.

Thus if we are to appraise the independent empirical content we need the type 3 comparisons as well. This role for type 3 comparisons <u>could</u>, in some cases, be taken over by a <u>conjunction</u> of type 1 and type 2 comparisons. However, if we consider the case where a theory is eliminated without it being falsified or shown to be improbable, we find that such appraisals can only be madeby means of a type 3 comparison; here they play a unique role. In this case we rely entirely on the excess corroboration of the new theory without reference to whether the old theory would be refuted by a type 2 comparison with respect to these novel statements. This is because the novelty of the statements is such that the question of their being a deductive consequence of, or a counter-instance of the old theory just does not arise. They are independent of the old theory. To the question of whether we could use just type 1 and type 3 comparisons for all our needs we must bring in the question of how rapidly we may achieve the goal of growth. We showed in Chapter Four that the pluralistic model of the deductive structure of theories which led to at least a hesitation in taking 'the facts' to be a basis of empirical comparison, indicated a slower rate of growth.

We saw also that if we are to interpret appraisals of growth as indications of increasing verisimilitude and reliability, we must also have recourse at least to type 3 comparisons and we gave methodological reasons for demanding the use of type 2 comparisons as well. Consideration of the Demarcation Principle of SMF led to what might be the overall conclusion of the essay: For the most rapid and empirically grounded progress our theories must be put into competition by as many methods and as often as possible.

Bibliography and References.

1 Lakatos: "Criticism and the Growth of Knowledge"
Edited by I. Lakatos and A. Musgrave
Cambridge University Press

2 Duhem: "The Aim and Structure of Physical Theory"
by Pierre Duhem, translated by P. Wiener
Princeton University Press 1954

3 Medawar: "The Art of the Soluble"
by P.B. Medawar
Methuen 1967

4 Poincare: "Les geometries non euclidiennes"
In: Revue des Sciences Pures et Appliquees, p 769 2 1891

5 Popper: "The Logic of Scientific Discovery"
by K.R.Popper (Third Edition)
Hutchinson 1968

6 Lakatos. op. cit. (1) p 116

7 Lakatos. op. cit. (1) p 116

8 Lakatos. op. cit. (1) p125

9 Kuhn: "The Structure of Scientific Revolutions"
by T.S.Kuhn
University of Chicago Press 1962

10 Popper op. cit. (5) p 137

11 Goodman: "Safety, Strength and Simplicity"
In: Philosophy of Science p 150 28 1961

12 Popper: "Conjectures and Refutations"
by K.R.Popper (Third Edition)
Routledge and Kegan Paul

13 Popper. op. cit. (5) p 137

14 Popper. op. cit. (12) preface

15 Wittgenstein: "Tractatus Logico-Philosophicus" by L. Wittgenstein, translated by D.F. Pears and B.F. McGuinnes. Routledge and Kegan Paul 1960

16 Jeffreys: "Scientific Inference" by H. Jeffreys Cambridge University Press 1937

17 Ackermann: "Inductive Simplicity" In: Philosophy of Science p 152 28 1961

18 Weyl: "Philosophy of Mathematics and Natural Science" by H. Weyl Princeton University Press 1949

19 Popper. op. cit. (12) p 223

20 Lakatos: "Changes in the Problem of Inductive Logic" In: The Problem of Inductive Logic; Proceedings of the International Colloquium in the Philosophy of Science, London 1965 North Holland Publishing Company 1968

21 Popper: "The Aims of Science" In: Ratio p 24 1 1957

22 Freud: "History of the Psychoanalytic Movement" In: Complete Psychological Works of Sigmud Freud Volume XIV p 59

23 Lakatos. op. cit. (1) p 128

24 Lakatos. op. cit. (1) p 128

25 Crookes: "The Study of Chemical Composition. An Account of its Method and Historical Development. by I. Freund Cambridge University Press 1904

26 Peirce: "Collected Papers" Volume 1. Cambridge University Press