CONCEPTS IN PHILOSOPHY OF SCIENCE

KEVIN McFARLANE

THE COMMONSENSE VIEW OF SCIENCE

Science is highly esteemed in modern times. The naming of some claim or line of reasoning or piece of research "scientific" is intended to imply a certain kind of merit or reliability. It is a widely held belief that there is something special about science and its methods.

> Scientific knowledge is proven knowledge. Scientific theories are derived in some rigorous way from the facts of experience acquired by observation and experiment. Science is based on what we can see and hear and touch, etc.. Personal opinion or preferences and speculative imaginings have no place in science. Science is objective. Scientific knowledge is reliable knowledge because it is objectively proven knowledge. [1]

The foregoing quotation sums up the popular commonsense view of what scientific knowledge is. The empirical method of physics is commonly regarded as the collection of facts by observation and experiment and the subsequent derivation of laws and theories from these facts by some sort of logical procedure. But modern philosophy of science, in particular the Popperian School, has pinpointed deep-seated difficulties in the idea that science rests on a sure foundation acquired through observation and experiment and with the idea that there is an inference procedure which allows us to derive scientific theories in a reliable way. There is no method which can prove scientific theories true or even probably true. Also, what is not often realised by people encountering Popperian ideas for the first time is that there is no method which enables scientific theories to be conclusively disproved either. Although this appears to imply complete scepticism it does not, as will be shown later.

WHAT PHILOSOPHY OF SCIENCE IS

Contemporary philosophy of science is not concerned with providing scientists with a rule book for solving their problems. How scientists arrive at their theories may be a subject-matter for the psychology and sociology of discovery. But modern methodologies or "logics" of discovery consist of a set of rules for the *appraisal* of ready, articulated theories. They also serve as theories of scientific rationality and attempt to provide explanations of the growth of objective scientific knowledge. The foregoing description is really only applicable to the rationalist wing of philosophy of science. Rationalists believe that there are criteria by which the relative merits of rival theories can be assessed independently of the beliefs, attitudes or other subjective states of individuals. Relativists believe that there is no objective way of distinguishing between theories and that the status of a theory will vary from individual to individual or from community to community.

The most influential group in contemporary philosophy of science is the Popperian School whose principal members I take to be Karl Popper, Imre Lakatos, Thomas Kuhn and Paul Feyerabend. Popper and Lakatos are rationalists. They believe that scientific theories can be objectively graded in terms of verisimilitude (closeness to the truth).[2] Kuhn and Feyerabend are relativists. I have included them in the Popperian School because they are dissenting Popperians. Kuhnian and Feyerabendian[3] positions are arrived at as two possible consequences of recognising that Popper's methodology is inadequate when tested against the history of science. A non-relativistic response to this fact is to develop an improved methodology. This is what Lakatos accomplished and it is with his philosophy that this essay is ultimately concerned. However, it will be helpful if some simpler philosophical positions are discussed first.

INDUCTIVISM

Inductivist theories of science assert that science starts with observations and that scientific theories are established from observation statements by some sort of generalisation procedure. Statements about the state of the world or some part of it can be established as true by an unprejudiced observer's use of his senses. Such statements can be termed "observation statements". Examples of observation statements are "Mr Smith struck his wife" or "That stick, partially immersed in water, appears bent". In other words, these statements refer to a particular occurrence or state of affairs at a particular place at a particular time. Scientific laws and theories, however, consist of general, *universal* statements. Examples are: "Planets move in ellipses around their sun" or "Acids turn litmus paper red". In other words, these statements make *unrestrictedly general* claims concerning all events of a particular kind at all places and at all times.

According to the inductivist these universal statements can be established by generalising from a finite set of observation statements provided that certain conditions are met. These are: (1) The number of observation statements forming the basis of a generalisation must be large, (2) The observations must be repeated under a wide variety of conditions, (3) No accepted observation statement should conflict with the derived universal law.

The first condition is intuitively obvious. We would not wish to conclude, for example, that all metals expand when heated on the basis of just one observation of a particular metal. However, if we only made a large number of observations of the same type, e.g. of one particular metal being heated in a particular way, we should be wary of making the generalisation also. This is why condition (2) is necessary. Various types of metal should be heated, at low temperatures, high temperatures, low pressures, high pressures, etc.. Condition (3) is obvious. Clearly any metal which is observed not to expand on heating would invalidate the universal statement.

The inductivist position can be summed up by means of the *principle of induction* as follows: "If a large number of A's have been observed under a wide variety of conditions, and if all those observed A's without exception possessed the property B then all As have the property B."

So far we have only examined how, according to the inductivist, scientific knowledge (consisting of observation statements and universal laws) is derived. To complete the picture it is necessary

Scientific Notes No. 5

ISSN 0267-7067 ISBN 0 948317 95 7

An occasional publication of the Libertarian Alliance,
25 Chapter Chambers, Esterbrooke Street, London SW1P 4NN
www.libertarian.co.uk email: admin@libertarian.co.uk

Kevin McFarlane is an engineer in the oil industry, designing and analysing offshore structures.

to consider how science arrives at predictions and explanations. This is accomplished by means of deductive inferences, i.e. logic. Given a verified observation statement and an inductively established universal law it is possible to predict or explain some phenomenon. For example, the following sequence of sentences constitutes a logically valid deduction. (1) Fairly pure water freezes at about zero degrees centigrade (if given sufficient time). (2) My car radiator contains fairly pure water. (3) If the temperature falls below zero degrees centigrade, the water in my car radiator will freeze (if given sufficient time). Statements (1) and (2) are the premises and (3) is the conclusion. The key feature of a logically valid deduction is that *if* the premises are true *then* the conclusions must be true. But the validity of the deduction itself is independent of the truth of the statements serving as the premises.

Of course, in practice deductive inferences can be much more complicated than the example described above. Typically, we could be faced with a large set of premises from which we would deduce the corresponding set of conclusions. In short, scientific inference can be construed as the derivation of predictions and explanations from initial conditions, laws and theories. This part of the picture is unproblematic and quite in accordance with other philosophies of science because it refers, essentially, to the *application* of knowledge already acquired. But the chief concern of the philosopher is the legitimacy or otherwise of that knowledge which is applied. The adequacy of the inductivist philosophy of science hinges ultimately on the reliability of observation statements and the legitimacy of the inductive procedure. These hypotheses will be questioned in the next two sections.

THE PROBLEM OF INDUCTION

The principle of induction referred to earlier was first seriously examined by David Hume[4] in the mid-eighteenth century. He demonstrated that the generalisation from singular observation statements to universal laws cannot be justified either by appealing to logic or experience. For example, no matter how many observations of black ravens have been made we can never justify the generalisation that all ravens are black. The next raven may always turn out to be white. This is obviously true regardless of whether the generalisation criteria of the principle of induction have been satisfied or not.

Another way of trying to justify the principle of induction is by appealing to the fact that induction has worked successfully many times in the past. For example, the laws of optics derived by induction from the results of laboratory experiments have been used frequently in the design of optical instruments which have functioned satisfactorily. But this argument is circular. It is saying in effect: the principle of induction worked successfully on occasion A; the principle of induction worked successfully on occasion B, etc.; therefore the principle of induction always works. But this employs the very kind of inductive argument the validity of which is supposed to be in need of justification. Incidentally, it could be asked here how one might attempt to justify the validity of logic without resorting to logic. The answer is that I don't think it can be done. But this is not a problem here. Logic is merely a set of rules for the tautological transformations of statements. To deny the validity of logic would yield the unwelcome conclusion that every imaginable statement is true, a conclusion which no sane person would accept.[5]

To avoid the difficulties just indicated one could reformulate the principle of induction like this: "If a large number of As have been observed under a wide variety of conditions, and if all these observed As without exception have possessed the property B, then all As probably possess the property B." But this is unsatisfactory also. Intuitively it does seem that the more observational support a universal law receives the more probable that it is true. But, as Chalmers indicates:

> ... any observational evidence will consist of a finite number of observation statements, whereas a universal statement makes claims about an infinite number of possible situations. The probability of the universal generalisation being true is thus a finite number divided by an infinite number, which remains zero however much the finite number of observation statements constituting the evidence is increased.[6]

This conclusion was established mainly by Popper's persistent efforts. It should be noted that the invalidity of inductive generalisation implies that scientific theories are unprovable.

In addition to the unavoidable logical difficulties described above it is possible to criticise the vagueness of the assumptions "a large number of observations" and "a wide variety of circumstances". How large is a "large number"? But also, is a large number necessary? Most people would not need to put their hands in a fire several times to conclude that fire burns. Also, most people do not need more than one atomic bomb to have been dropped to conclude that atomic bombs cause widespread death and destruction. In a different vein, a "wide variety of circumstances" must in practice mean an infinite number unless it is decided that some variations are superfluous. But the significant variations can only be distinguished from the superfluous ones by our *theoretical knowledge* of the situation. This implies that theory plays a vital role *prior* to observation and this cannot be admitted by the most naive inductivist. However, some inductivists are prepared to concede that theory *can* precede observation but they still maintain that the *adequacy* of a theory can only be assessed by recourse to some sort of principle of induction. But there are additional problems in the inductivist programme which I shall describe below.

THE THEORY DEPENDENCE OF OBSERVATION

A central tenet of the inductivist account of science is that observation yields a secure basis from which knowledge can be derived. But in the case of visual perception, for example, our perceptual experiences do not have a simple one-to-one correspondence with our retinal images. In other words, two normal observers viewing the same object from the same place under the same physical circumstances do not necessarily have identical visual experiences, even though the images on their respective retinas may be virtually identical. The experience of optical illusions, for example, confirms this. In our culture the practice of depicting three- dimensional objects by two-dimensional perspective drawings enables us to see things like staircases as staircases. But experiments on members of certain African tribes indicate that they do not see our three-dimensional drawings but merely arrays of two dimensional lines.[7] What an observer sees, that is, his visual experience, depends in part on his past experience, knowledge and expectations. More conventional scientific observations such as those obtained from a telescope or from a medical X-ray machine require prior *theoretical knowledge*. For example, a medical student viewing X-ray pictures sees hardly anything that the experienced observer does. Or, consider trying to establish that a short white stick found near a blackboard is a piece of chalk. It may just *look like* a piece of chalk. If experiments are set up to demonstrate that it is a piece of chalk this entails recourse to *further* observations and also presupposes the operation of certain theoretical laws (i.e. relating cause and effect). In principle the chain of experiments in response to the objections of doubting observers could extend indefinitely. But, in practice, agreement is usually reached at some point along the chain, since the different observers will share a number of common theoretical assumptions.

In addition to the comments raised at the end of the last section it is clear from the foregoing that the observation process is problematic.

INTRODUCTION TO FALSIFICATIONISM

A philosophy of science which accepts the theoretical status of observation statements as well as the hypothetical character of all scientific laws is falsificationism. If it is temporarily assumed that observation statements are secure then it still remains true that universal laws are unprovable. However, the falsificationist points to a logical asymmetry between verification and falsification. For example, consider the statement: "All ravens are black." Although no number of observations of black ravens enables us to prove, i.e.

verify, the validity of this law, a single observation of a non-black raven will falsify it. Therefore, although universal laws are empirically unprovable, they *are* empirically falsifiable. Thus, at the level of *logic*, falsificationism regards science as a process of conjectures and refuations. Theories are proposed as solutions to existing problems in the body of science. These theories are then rigorously tested. Those which fail the tests are refuted while those which pass the tests are "verified", that is, tentatively accepted until further tests cause us to revise our verdicts. While it cannot be said of any theory that it is true, it can be said that a particular theory is the best available, that it has stood up to all its tests.

Since scientific theories need to be tested, and given the logical asymmetry between verification and falsification, the falsificationist proposes falsifiability as a *criterion of demarcation* between the scientific and the non-scientific. It should be noted that "non-scientific" does not mean "irrational". In this context it means *empirically unfalsifiable*. It is worth remarking here that critics of Popper have often taken his falsifiability criterion to be a means of distinguishing between "superior" and "inferior" knowledge. For example, the logical positivists of the Vienna Circle[8] thought that Popper had proposed an alternative *criterion of meaning* for demarcating sensible statements from nonsensical ones.

Another reason for proposing the falsifiability criterion is that science, unlike, say, logic or mathematics, is concerned with obtaining knowledge about the external world. Falsifiable statements possess what is referred to as *informative content* whereas unfalsifiable ones do not. Now an hypothesis is falsifiable if there exists a logically possible observation statement which is inconsistent with it, that is, which, if established as true, would falsify the hypothesis. We have already seen that the statement: "All ravens are black" is falsifiable. But consider the following statements. "Either it is raining or it is not raining." "Luck is possible in sporting speculation." It is clear that no logically possible observation statement could refute either of these two statements. Therefore, they tell us nothing about the world. But a falsifiable statement asserts that the world does, in fact, behave in a certain way, thereby *ruling out* ways in which it could logically behave but does not. The law: "All planets move in ellipses around the sun" is scientific because it claims that planets, in fact, move in ellipses and *rules out* orbits that are square or oval. Just because the law makes definite claims about planetary orbits, it has informative content and is falsifiable.

Thus far, for simplicity, I have described falsificationism in a manner which assumes that observation statements can be unequivocally decided. This, as we have seen, is false. I intend to develop the falsificationist programme further by first considering what Lakatos has termed naive falsificationism.[9]

NAIVE FALSIFICATIONISM

In the previous section it was assumed that what is true at the level of logic can be applied unproblematically to objective reality. The falsificationism defined thus far, which we shall term dogmatic falsificationism, is, like inductivism, a justificationist theory of knowledge. It holds that the truth value of factual propositions can be indubitably decided. This implies that there is a natural borderline between observation statements and theoretical statements. This, of course, corresponds to the commonsense view but is, in fact, mistaken. Nevertheless we cannot avoid using commonsense terminology in normal, "low-level", everyday discourse. In discussing epistemology, though, we need to know what is really going on.

The application of logic to reality, however, is by no means a simple matter, due, among other things, to the theory-dependence of observation. Another difficulty for dogmatic falsificationism can be cited here. Scientific theories forbid an event occurring on condition that no other relevant factor has any influence on it. Therefore, a falsifying statement really refutes a theory plus a universal non-existence statement (termed a *ceteris paribus* clause) saying that no other relevant cause is at work anywhere in the Universe. It is always possible to save a theory by claiming that there is some factor at work influencing the result and then proposing an experiment to test for this factor. Moreover, *ceteris paribus* clauses are normally the rule in science rather than the exception.

Thus, in realistic test situations we are confronted with the theory under test, the initial conditions (in this case, comprising falsifying statement(s)) and a *ceteris paribus* clause. Given a falsifying statement(s), there are three possibilities. (1) The falsifying statement(s) is true and the theory is refuted. (2) The initial conditions have been established incorrectly. A mistake has been made in the observations; or the generally accepted background theory, in the light of which the observations have been established as "true", is false. (3) The *ceteris paribus* clause is false and some other factor is at work. So the "falsification" of a theory cannot be as straightforward a matter as dogmatic falsificationism would have us believe.

However, it is possible to overcome the foregoing difficulties by proposing naive [methodological] falsificationism. The main characteristic of this epistemology is the introduction of *methodological decisions* which supplant the *assumptions* of dogmatic falsificationism. At this point it is helpful to introduce some abbreviations. We shall designate dogmatic falsificationism as DF, naive falsificationism as NF (a somewhat unfortunate choice!) and the sophisticated falsificationism to be described in the next section as SF. ("NF" in Britain is short for "National Front", a British national socialist group. Ed.)

NF consists of the following ingredients. It makes unfalsifiable by *fiat* some singular statements distinguished by the fact that there exists a relevant technique such that anyone who has learned it will be able to "decide" that the statement is "acceptable" or "observational". These observation statements will either be *potential falsifiers* of a particular theory or *potential corroborators*, i.e., they will confirm the theory. "Confirm", in this context, means that a theory passes a test. It does not mean "prove" since it is impossible to prove a scientific theory. The theories which are used to fortify the observation statements comprise, by decision, the unproblematic background knowledge, which is accepted as unproblematic while the problematic theory is being tested. For example, observations through a radio-telescope would be corroborated on the basis of the [unproblematic] theory of radio-optics.

Essentially, the problematic and the unproblematic background knowledge must be decided by agreement. If agreement is not reached we keep testing until it *is* reached. At this point, or at various "agreed" points along the chain of testing, everybody is obliged to abide by the results - at least under the NF regime. The conventionalism of this procedure ultimately has implications which undermine the NF methodology as we shall see.

The following example will illustrate how NF tackles the problems alluded to earlier concerning DF. Consider the theory of, say, Newtonian mechanics. Then let N3 comprise the treble conjunction of Newton's theory, the initial conditions (which represent, say, Mercury's anomalous planetary orbit - a potential falsifier of Newton's theory), and a *ceteris paribus* clause. The initial conditions are tested severely, i.e. the potential falsifier is corroborated and relegated into the unproblematic background knowledge. Thus what is now being refuted is N2, comprising Newton's theory plus the *ceteris paribus* clause. It is decided to test the *ceteris paribus* clause severely. To do this it is necessary to specify certain potential influencing factors and to test for their effects. If the *ceteris paribus* clause is corroborated, the initial conditions now constitute a refutation of N1, Newton's specific theory.

It is clear that the decision to accept a *ceteris paribus* clause is very risky since any number of factors may be influencing the initial conditions and we cannot possibly specify all of them. In fact, those that we *do* specify will be dependent on our current, unproblematic, theoretical knowledge. Nevertheless, at some stage in the chain of testing we *do* need to accept a *ceteris paribus* clause otherwise there will be no way of eliminating any theories and the growth of science will be just the growth of chaos. I will say more about this later.

Another point can be made here concerning the previous example. Prior to the acceptance of the *ceteris paribus* clause the initial conditions merely constitute an "anomaly" in relation to N1. But after its acceptance they constitute a *falsification* of N1. To quote Lakatos:

> We may call an event described by a statement A an "anomaly in relation to a theory T" if A is a potential falsifier of the conjunction of T and a *ceteris paribus* clause but it becomes a potential falsifier of T itself after having decided to relegate the *ceteris paribus* clause into "unproblematic background knowledge".[10]

In the process of testing a theory it is important to note that the falsifying or corroborating statements must be "inter-subjectively testable". This means that it must be possible for more than one experimenter to reproduce the results of an experiment (within the limits of experimental error). Precisely how many times an experiment should be repeated is a matter of convention. But clearly if an experiment is only performed once it cannot properly be accepted by the scientific community. Science cannot operate with stray or "occult" phenomena.[11] This is commonsense (which, in this case, is correct).

Considering the truth or falsity of a theory, NF separates "rejection" and "disproof" where DF had conflated them. For NF, rejection is a methodological decision which arises as a result of a procedure similar to the one given in the foregoing example. Due to the inherent risks of the process it is quite possible to accept a false theory and reject a true one. This must seem quite paradoxical to the dogmatic falsificationist (and to anyone else, in fact).

Finally, under NF, it becomes possible to accept probabilistic theories. For DF, probabilistic theories should strictly be rejected as unscientific because they cannot be falsified. But for NF, probabilistic theories can be falsified by specifying rejection rules which render statistically interpreted evidence "inconsistent" with a particular probabilistic theory.

The conventionalist procedure I have just outlined may seem somewhat artificial and arbitrary. Scientists are not consciously aware of behaving in accordance with a particular methodology such as that of naive falsificationism. But what is being described here is a sort of "rational reconstruction" of the growth of scientific knowledge. The adequacy of the conventionalist methodology should be gauged by reference to the purpose for which the methodology is chosen. In this case the purpose is the maximisation of consistent, empirical informative content (i.e. the attempt to discover as much as possible about the behaviour of reality). Yet considered as an attempt to provide a rational reconstruction of the history of science naive falsificationism must be regarded as inadequate. A theory of scientific rationality should be able to reconstruct the generally agreed best theories and the steps leading to those theories as rational. That is, if too many of the actions of individual scientists in their development of the generally agreed best theories are irrational in the light of a particular methodology then that methodology should be rejected as inadequate. For example, Lakatos writes:

> While there has been little agreement concerning a *universal* criterion of the scientific character of theories, there has been considerable agreement over the last two centuries concerning *single* achievements. While there has been no *general* agreement concerning a theory of scientific rationality, there has been considerable agreement concerning whether a particular single step in the game was scientific or crankish ... A general definition of science thus must reconstruct the acknowledgedly best gambits as "scientific": if it fails to do so it has to be rejected.[12]

One implication of NF methodology is the periodic recurrence of so-called "crucial experiments" which allegedly decide between two or more competing theories. For example, Mercury's anomalous perihelion (a peculiarity of its planetary orbit) could have been said to be a decisive refutation of Newtonian gravitational theory. But, as Lakatos writes:

> eighty-five years elapsed between the acceptance of the perihelion of Mercury as an anomaly and its acceptance as a falsification of Newton's theory, in spite of the fact that the *ceteris paribus* clause was reasonably well corroborated.[13]

Also, stubborn theoreticians frequently challenge experimental verdicts and have them reversed. This is irrational behaviour from the point of view of NF. For example:

> Galileo and his disciples accepted Copernican heliocentric mechanics in spite of the abundant evidence *against* the rotation of the Earth; or Bohr and his disciples accepted a theory of light emission in spite of the fact that it ran counter to Maxwell's well corroborated theory.[14]

Considering the fact that such "irrational" behaviour can be shown (as will be seen subsequently) to be very fruitful, this suggests that the methodology of NF, while being logically impeccable, is too rash. It requires "refutations" to kill theories too quickly. Moreover, it tends (but not exclusively) to regard testing as largely an isolated fight between a theory and an experiment in which the only interesting outcome is [methodologically] conclusive falsification. But, in fact, tests are usually:

> three-cornered fights between rival theories and experiment and some of the most interesting experiments result, *prima facie*, in confirmation rather than falsification.[15]

The solution to these difficulties is to improve the conventionalist elements of NF to allow greater scope for "irrational" maneouvres, i.e. to make those "irrational" maneouvres "rational" with respect to an improved methodology. Such an improvement can be provided by sophisticated falsificationism (SF). It is to this which we shall now turn.

SOPHISTICATED FALSIFICATIONISM

It may seem that the proposal to develop an improved methodology in order to make irrational behaviour "rational" is something of a fiddle. A critic might remark: "This fellow, the naive falsificationist, has proposed a wonderfully elegant philosophy of science, logically watertight and able to provide a rationale for the growth of scientific knowledge. Unfortunately some of our best theories, which most of the scientific community have accepted as such, seem to have been arrived at by appallingly irrational behaviour if one consults the history of science. But instead of this falsificationist admitting the inadequacy of this theory and concluding that the growth of science is irrational he prefers to concoct some *ad hoc* assumptions to rationalise the irrational."

But, in actual fact, the proposal to improve NF is not some rearguard action. It is merely a positive, constructive response to criticism. It seems fairly obvious to me and, no doubt, to many others that Einsteinian mechanics is, objectively, a theory of greater informative content and is more satisfactory in its performance, say, in high energy physics than Newtonian mechanics. If the steps leading away from Newton and towards Einstein appear, historically, to be irrational it is nevertheless clear that such irrational steps have led to a better theory. Therefore, it ought to be possible to provide a rationale for these irrational steps. It should be possible to demonstrate how such irrational behaviour can bear fruit. It can then be alleged that this behaviour is, in fact, not necesssarily irrational in the context of an improved methodology. All this will become clearer once the principles of sophisiticated falsificationism and, subsequently of the methodology of scientific research programmes are elaborated.

The essential innovation introduced by SF is the recognition of the continuity of scientific knowledge and, therefore, of the fact that the unit of appraisal is not an isolated theory or theories but a *series* of theories. The conception of SF is somewhat technical and it will be convenient to summarise the basic ingredients in a number of sections with some practical illustrations appearing at appropriate intervals.

(1) A theory is scientific if it has excess empirical content over its predecessor, that is, if it leads to novel facts. This assertion can be broken down into three stages.

(i) A theory T' has excess empirical content over a theory T.

(ii) T' explains the previous success of T, i.e. all the unrefuted content of T is included within T'.

(iii) Some of the excess content of T' is corroborated.

SF enables theories to be saved from refutation by the postulation of auxiliary hypotheses. But such hypotheses must increase the informative content of the theory or series of theories. Mere linguistic adjustments are ruled out as *ad hoc*. A simple example of unacceptable and acceptable theory modification is provided by Chalmers.[16] Consider the claim that "bread nourishes". This simple assertion ought really to be spelled out in more detail but for this simple example we shall assume that its meaning is sufficiently clear. Now suppose that this statement is "falsified" by the discovery of some people who have become ill and died from consuming bread. If the theory is modified to the statement: "all bread nourishes with the exception of the particular batch of bread consumed by this particular group of people" then this modified theory cannot be tested in any way that was not also a test of the original theory. The consuming of any bread by any human constitutes a test of the original theory, whereas tests of the modified theory are restricted to the consuming of bread other than the batch which did not nourish. Thus, this modification is *ad hoc* and is rejected by the sophisticated falsificationist. Now suppose the theory is modified in the following manner: "All bread nourishes except bread made from wheat contaminated by a particular kind of fungus (followed by a specification of the fungus and some of its characteristics)." This modified theory is not *ad hoc* because it leads to new tests. It is *independently testable*. A possible test would be testing the wheat from which the poisonous bread was made for the presence of the fungus. Chalmers lists further tests. All of these tests would result in the falsification of the modified hypothesis. If they do not then something new will have been learnt and progress will have been made. The original theory will have been modified in a *progressive* manner. There are many examples from the history of science which could be used to illustrate *ad hoc* and progressive auxiliary hypotheses. The reader is referred to Chalmers and Lakatos.[17]

(2) A point which has already been mentioned. Theory appraisal passes to a series of theories (or modifications) rather than to an individual theory.

(3) Falsifying theories are frequently proposed *after* the counterevidence. Until then the counterevidence is just an "anomaly".

For example, recall the case of Mercury's perihelion which was not accepted as a falsification until eighty-five years after its discovery, i.e. until the emergence of Einstein's general theory of relativity - which was able to explain the previous success of Newton's theory, account for the anomalous behaviour of Mercury and also predict some new phenomena (some of which were corroborated). There are, in fact, anomalies in relativity theory as well, as there are in most theories. But these merely constitute research problems. Since, at present, there is no adequate rival to general relativity, the anomalies are not considered refutations.

(4) It is not just falsification which can be decisive. Cases of *verified* excess empirical content can be decisive too.

At this point it is worth clarifying the assertions made under (1) above. The concept "empirical content" is derived from the concept "informative content".[18] The informative content of a theory is the set of statements which are incompatible with the theory. (That is, theories tell us more the more they prohibit or exclude). The empirical content of a theory is the class of its *potential falsifiers*. That is, the empirical content of a theory refers to those aspects of it which are *empirically* testable. If a theory T' has excess empirical content over a theory T this implies that T' has some testable consequences which are not also consequences of T. For T' to be *entirely* [theoretically] progressive with respect to T it should be able to explain all of the content of T as an approximation to T'. For example, the special theory of relativity is able to explain how the simple Newtonian schema for the calculation of relative velocities between objects *has* worked in the past and *does* work now for objects moving at low velocities compared with that of light.

It is possible for theory T' to have excess empirical content over T and *vice versa*. In such a case one might have to employ both theories. It is also possible for T' to be *entirely* theoretically progressive with respect to T but for none of this excess theoretical content to be verified. In such a case T' is rejected.

The stress that SF lays on rival, overlapping theories constitutes what Lakatos and Popper term a problem shift. For example, appraisal has passed from an isolated theory to a series of theories. We always have to consider the historical problem-situation in which a particular theory or auxiliary hypothesis is proposed. Significant advances in science can be represented by *either* verifications *or* falsificaitons. We learn the most when either a *bold*, daring conjecture is confirmed or when a *cautious* conjecture (i.e. a piece of generally accepted and corroborated knowledge) is refuted. If Einstein's prediction that light rays should be bent by the sun's gravitational field had been refuted this would not have caused much of a stir (but it *would do* if this happened today rather than in 1917). Similarly if I verify a prediction of Newtonian mechanics or Maxwellian electromagnetism today, I contribute nothing of importance to scientific knowledge. But if I falsify Maxwell's theory, in a field of application in which it is supposed to be satisfactory, then this would be significant. (We would not, however, *necessarily* reject Maxwell's theory unless some new theory were able to explain what was going on.)

One implication of SF is that the so-called "crucial experiment" is something which can only be recognised *with hindsight*. An experiment is only crucial in the context of a new theory. This will become even clearer when the methodology of scientific research programmes is elaborated.

Under SF, because falsifications need not be accepted as such immediately this enables scientists to treat them as research problems. There is also greater scope for the theoretician to challenge experimental verdicts by questioning the interpretive theory in the light of which the observations were made. Nevertheless, SF still regards *ad hoc* modifications to theories as unscientific and also still suggests a regime in which theories are killed off too quickly. One can, in fact, improve SF by adding on a few more liberalising rules but it is best to change to a different formulation which is more-or-less implicit in SF, i.e. the methodology of scientific research programmes (MSRP).

THE METHODOLOGY OF SCIENTIFIC RESEARCH PROGRAMMES

MSRP extends the basic ideas of conventionalism introduced by methodological falsificationism to create a more liberal regime in which to appraise the actions of individual scientists and their theories.

The shift of the unit of appraisal from isolated theories to series of theories instituted by SF points to the more coherent concept of a *research programme*. The Lakatosian conception introduces some rather strange terminology, the meaning of which will become clearer shortly. As an introduction I will quote from Chalmers:

> A Lakatosian research programme is a structure that provides guidance for future research in both a positive and a negative way. The *negative heuristic* of a programme involves the stipulation that the basic assumptions underlying the programme, its *hard core*, must not be rejected or modified. It is protected from falsification by a *protective belt* of auxiliary hypotheses, initial conditions, etc.. The *positive heuristic* is comprised of rough guidelines indicating how the research programme might be developed. Such development will involve supplementing the hard core with additional assumptions in an attempt to account for previously known phe-

> nomena and to predict novel phenomena. Research programmes will be *progressive* or *degenerating* depending on whether they succeed in leading or whether they persistently fail to lead to the discovery of novel phenomena.[19]

Let us examine these concepts in more detail.

(1) *The Negative Heuristic*

The negative heuristic or "hard core" of a research programme is a set of univeral statements rendered irrefutable by the *methodological decision* of its proponents. Refutations are forbidden to be directed at this hard core. Instead the hard core is surrounded by a protective belt of auxiliary hypotheses which must bear the brunt of any refutations. The existence of the refutations poses a continual challenge to the theoreticians in modifying (sophisticating) the protective belt in order to account for them.

Newtonian mechanics is an example of a research programme in which the hard core consists of the three laws of motion plus the law of gravitation. The protective belt consists of observational theories, initial conditions and mathematical techniques, e.g. for solving differential equations. A progressive research programmes is one which displays at least an intermittently *empirical* progressive problemshift (in the manner of sophisticated falsificationism). A *theoretically* progressive problemshift is one in which some novel fact is predicted (or some previously unexplained phenomenon or "anomaly" is accounted for) in a *content-increasing* manner. A research programme, as a whole is, empirically progressive if every now and then some of the new theoretical informative content is corroborated. At any particular step corroboration need not be required immediately. To quote Lakatos:

> We do not demand that each step produce immediately an observed new fact. Our term *'intermittently'* gives sufficient *rational* scope for dogmatic adherence to a programme in face of *prima facie* "refutations".[20]

So long as the corroborated empirical content of the protective belt of auxiliary hypotheses increases we may rationally decide not to allow "refutations" to transmit falsity to the hard core. We need not reject a research programme *merely* because it contains anomalies. Virtually all such programmes *do* contain and *have* contained anomalies. How research programmes can be eliminated, if at all, will be discussed later.

(2) *The Positive Heuristic*

The positive heuristic of a research programme is a long-term research policy which anticipates and digests "refutations", which are regarded as "anomalies" until the programme either degenerates or is superseded.

> The positive heuristic consists of a partially articulated set of suggestions or hints on how to change, develop the "refutable variants" of the research programme, how to modify, sophisticate, the "refutable" protective belt.[21]

The positive heuristic of a programme forges ahead virtually regardless of refutations. So long as the programme is progressing refutations are irrelevant. Progress takes shape as a series of ever more complicated physical models. Usually it is known that these models are inadequate both theoretically and empirically. (For example, they may contain contradictions). In fact, it may be that early on in a research programme the development is entirely theoretical and the data are not even looked at.

The emphasis that MSRP places on corroborated novel content suggests that it is the *verifications* rather than the refutations which provide the heuristic driving-power - though the verification of the (n+1)th version of a programme is, of course, the refutation of the nth.

A degenerating research programme is one in which either none of its theoretically progressive informative content can get corroborated or one in which refutations can *only* be dealt with *ad hoc* (in a content-decreasing manner). Lakatos has provided a number of practical examples from the history of science to illustrate certain aspects of MSRP.

For example, Newtonian mechanics was (and still is) one of the most successful research programmes ever. The history of its development provides beautiful examples of the rationale of MSRP. MSRP is able to reconstruct a much larger part of the history of science as rational than can other philosophies of science. Nevertheless not all of the history of science is rational since human beings are not perfectly rational creatures (in the sense of behaving honestly, reasoning correctly, etc.).

Incidentally, although Newtonian mechanics has been superseded by Einsteinian mechanics, it is, nevertheless, still a progressive research programme within its *legitimate field of application*, so to speak (i.e. in the domain of low velocities and weak gravitational fields). The Newtonian programme is still forging ahead in the form of modern applied mechanics, e.g. engineering research.

(3) *Competition Between Research Programmes*

Of an individual research programme it is possible to say, objectively whether it is progressive or degenerating. In a degenerating phase of a particular programme the theoreticians' attention will tend to turn to the hard core and this will be considered open to refutation. *However*, no-one can be sure that a research programme that is currently degenerating may not later turn into a progressive one again, due to the creativity of some scientists. This seems to pose a problem which I shall discuss in more detail later.

Thus far we have only analysed the characteristics of *individual* research programmes. In reality there is usually competition between programmes. So the important question is how to compare rival research programmes. Essentially, a solution to this matter has already been provided within the methodology of sophisticated falsificationism. A research programme R' supersedes a research programme R if and only if R' has excess corroborated content over R but not *vice-versa.* Thus, if R' is able to explain everything that R explains and is able to predict novel facts or account for previously unexplained ones, then R' supersedes R from an empirically informative point of view. However, from a *technological* point of view R can still be used, might even be preferred (because it is simpler) and could still be progressive within its domain of application (e.g. Newtonian mechanics).

One consequence of all this is that a so-called "crucial experiment" can only be considered as such after *long* hindsight - *not* at the time of the experiment. It is only the historian and philosopher of science who can bestow the honorific title of crucial experiment after comparing the success of two rival research programmes. As Lakatos writes:

> When one research programme suffers defeat and is superseded by another one, we may - with long hindsight - call an experiment crucial if it turns out to have provided a spectacular corroborating instance for the victorious programme and a failure for the defeated one (in the sense that it was never explained "progressively" - or briefly, "explained" - within the defeated programme).[22]

THE SCOPE OF RATIONAL SCIENTIFIC ACTIVITY

In the original conception of falsificationism, Popper proposed falsifiability as a criterion of demarcation between science and nonscience. This criterion served two purposes: (1) to distinguish between scientific theories and pseudo-scientific theories, and (2) to distinguish between scientific *activity* and pseudo-scientific activity. Thus, individual scientists can be judged to be behaving scientifically or not according to Popper's philosophy. The consequence of this is that a great deal of scientific behaviour must be regarded as irrational as we have seen. Lakatos' MSRP constituted a problemshift as far as the proposals of criteria of demarcation were concerned. In the Lakatosian scheme there is no clearcut criterion of demarcation. The demarcation is one of degree - between progressive and degenerating programmes or between research programmes and disconnected chains of conjectures and refutations, lacking in unity. Moreover, the appraisal passes fully to ready, articulated theories embedded in research programmes and the scientist, as actor, is free to do whatever he likes. The Lakatosian regime is much more liberal and "anarchistic" than the Popperian regime.

According to Lakatos, Popper conflates the two positions of naive falsificationism and sophisticated falsificationism. Thus, in some of his utterances he has demanded naively that:

> ... criteria of refutation have to be laid down beforehand: it must be agreed, which observable situations, if actually observed, mean that the theory is refuted.[23]

But Popper also maintains that some degree of dogmatism is fruitful, for without it we may never know the potential strength of a theory. However such a rationale cannot be satisfactorily accounted for in Popper's methodology.

MSRP is only an appraisal of science as a body of *objective knowledge*, with theories having objective consequences and interrelationships which are independent of the subjective states or behaviour of individual scientists. Therefore, although scientists, in their actions, are obviously not aware of MSRP, it is nevertheless possible to apply MSRP as an *appraisal* system. Although in discussing MSRP we have referred to "decisions" this is really just a figure of speech in order to facilitate understanding.

Concerning the behaviour of individual scientists MSRP permits *ad hoc* content-reducing strategems, the evasion of "refutations" or "anomalies" and the questioning and correction of experimental verdicts. To less sophisticated epistemologists such behaviour must lead to the conclusion that the game of science is irrational. But such a conclusion would be false. Since it is virtually impossible instantly to produce a theory which is fully consistent both internally (i.e. in terms of logical consistency) and externally (i.e. in terms of empirical corroboration) it *is* rational to regard refutations as anomalies. But a progressive research programme must be able to digest *some* of the anomalies. What is *not* rational is to *deny* that anomalies are anomalies. Thus it *is* rational to continue working on a degenerating programme but it *is not* rational to deny that it has a poor "public" record. (A degenerating programme may always stage a comeback.) This poses the question of when to stop working on a degenerating programme. Because Lakatos provides no criteria for determining when to abandon a degenerating research programme both Kuhn and Feyerabend have criticised MSRP as nothing more than a "verbal ornament".[24] But both Kuhn and Feyerabend miss the point. MSRP just appraises the objective state of knowledge. It does not tell the scientist which problems to work on. The issue is largely a matter of commonsense. If a degenerating research programme has no rival then clearly scientists have no choice but to work in that programme. (In practice, a research programme usually *will* have a rival or there will be competition between "sub-programmes" of the main research programme - for example, the competition between various analytical methods in engineering research.) If a research programme has a *superseding* rival then scientists can either keep working on the old programme within its valid domain of application or they can try to regenerate it in its "refuted" domain. However, if the superseding programme is forging ahead then, sociologically, scientists will tend to work on that programme and research funds will flow in that direction. The dissenting scientists are perfectly free to continue working on the degenerating programme if the research funds can be found. Chalmers has developed a theory of the rationale which governs the tendency of scientists to work on one programme or another.[25]

Both Kuhn and Feyerabend confuse the sociology and psychology of knowledge with the logic of knowledge - an error which is shared by Marxism also.

In this essay I have attempted to give a simplified account of modern philosophy of science based on the Popperian critical rationalist wing. Popper's most important achievement has been to show the fallibility of all theories purporting to make empirical claims about the world and, as a corollary, to emphasise the necessity for the critical analysis of all theories. Both the theories and their refutations are permanently open to question. In any particular research programme anomalies can be ignored (in the sense that they are not dealt with immediately) but it is always acknowledged that they *are* anomalies and are, therefore, problematic.

In a subsequent essay I intend to discuss the implications of Popperian/Lakatosian philosophy of science and its relationship, if any, to libertarian philosophy.

REFERENCES

1. A. F. Chalmers, *What Is This Thing Called Science?*, Open University Press, 1983.
2. The concept of verisimilitude is problematic. For a fuller discussion see reference 1, chapter 13.
3. See T. S. Kuhn, *The Structure of Scientific Revolutions*, Princeton University Press, 1970; and P. K. Feyerabend, *Against Method: Outline of an Anarchistic Theory of Knowledge*, New Left Books, London, 1975.
4. D. Hume, *Treatise on Human Nature*, Dent, London, 1939.
5. See K. R. Popper, "What is Dialectic?" in *Conjectures and Refutations*, Routledge and Kegan Paul, London, 1963.
6. See reference 1.
7. N. R. Hanson, *Patterns of Discovery*, Cambridge University Press, 1958.
8. The Vienna Circle was an intellectual gathering of Logical Positivists in the 1930s. Popper had a close association with this group but always as a critic. His *The Logic of Scientific Discovery* is, in fact, a quite extensive refutation of Logical Positivism.
9. I. Lakatos, *The Methodology of Scientific Research Programmes*, Philosophical Papers Volume 1, title essay, Cambridge University Press, 1980.
10. *Ibid.*
11. For example, the difficulty in dealing with the paranormal is not so much that the occurences violate well-corroborated scientific theories but that the occurrences are, in general, not repeatable - not amenable to controlled experimentation. It may be that what is required is a different methodology for comprehending such phenomena.
12. I. Lakatos, "The History of Science and Its Rational Reconstructions" in *The Methodology of Scientific Research Programmes*, Philosophical Papers Volume 1, Cambridge University Press, 1980.
13. See reference 9.
14. *Ibid.*
15. *Ibid.*
16. A .F. Chalmers. See reference 1, chapter 5.
17. See references 1 and 9.
18. For a fuller discussion of this terminology see K. R. Popper, *Unended Quest: An Intellectual Autobiography*, chapter 7, Fontana, 1976. For a highly detailed discussion see K. R. Popper, *The Logic of Scientific Discovery*, chapter 6, Hutchinson, London, 1959.
19. See reference 1, chapter 7.
20. See reference 9.
21. *Ibid.*
22. *Ibid.*
23. K. R. Popper, *Conjectures and Refutations*, Routledge and Kegan Paul, London, 1963.
24. T. S. Kuhn, "Reflections on my Critics" in I. Lakatos and A. Musgrave, *Criticism and The Growth of Knowledge*, Cambridge University Press, 1970. P. K. Feyerabend, "Consolations For The Specialist" in I. Lakatos and A. Musgrave (eds.), 1970.
25. See reference 1, chapter 11.

RECOMMENDED READING

As the best, most accessible introduction to the ideas discussed in this essay I recommend: A. F. Chalmers, *What Is This Thing Called Science?*, Cambridge University Press, 1983.