The Problem of Rational Theory-Choice

Howard Sankey
Department of History and Philosophy of Science
University of Melbourne

Introduction

The problem of rational theory-choice is the problem of whether choice of theory by a scientist may be objectively rational in the absence of an invariant scientific method. In this paper I offer a solution to the problem, but the solution I propose may come as something of a surprise. For I wish to argue that the work of the very authors who have put the rationality of such choice in question, Thomas Kuhn and Paul Feyerabend, contains all that is needed to solve the problem.

The problem of rational scientific theory-choice is a problem which was generated out of the clash between the two major twentieth century traditions in the philosophy of science. On the one hand, there is the empiricist tradition in the philosophy of science, which includes logical positivism and falsificationism, and which holds that science is governed by a single scientific method, invariant throughout history and the various branches of science. On the other hand, there is the more recent historical approach to philosophy of science, the main advocates of which have been Kuhn and Feyerabend, which takes the practice of science to vary with historical time-period, theoretical context, and scientific discipline. In contrast with the former, the latter allows that there may be variation in

scientific methodology throughout the history of science, and between different branches of science.

The question of whether scientists' choice of theory may be rational arose in the wake of widespread rejection of empiricist models of science during the 1950s and 1960s. According to the empiricist conception of science, rational acceptance of theories is governed by an invariant scientific method which is applicable throughout science. This method typically involves the use of observational data, as a basis on which to generalize to universal theories, whose logical connection with the data is either inductive or deductive in form. The principal reasons for abandoning such models have involved problems with the empirical basis of science due to the theory-dependence of observation and the underdetermination of theory by data. In addition, problems of both a historical and philosophical nature have raised serious doubts about the existence of a uniform scientific method characteristic of science throughout its history.

According to empiricist philosophy of science, the acceptance of a theory by a scientist is rationally justified provided the scientist's acceptance of the theory is certified by the scientific method. For example, if a theory has attained a high degree of confirmation based on empirical evidence which supports it, then acceptance of the theory is rationally justified. Alternatively, if a theory has been submitted to severe tests without being refuted, and no other theory is as well-tested, then it is rational to accept the theory. In either case, rational theory-acceptance is based on objective grounds, since both observation and logical inference provide epistemically

well-grounded procedures of inquiry, and because such procedures jointly constitute a neutral court of appeal to which all scientists have recourse. Theory-acceptance, therefore, need not be made on the basis of subjective matters of personal whim, bias or taste, since it may be based on methodological considerations, which are objective in the sense both of having properly epistemic import and being open to public scrutiny.

Advocates of the historical approach, by contrast, argue that there is no fixed scientific method, and that there is instead variation in the methodological standards employed by scientists in the evaluation of theories. Yet, in the absence of a fixed scientific method, the historical school is unable to account for the rationality and objectivity of scientific theory-acceptance in the manner of empiricist philosophy of science.

This leads to the problem of rational theory-choice. For if there is no fixed scientific method, then it is unclear how the choice of one theory over another can be either rational or objective. This is particularly the case if the advocates of one theory endorse one set of methodological standards, and the advocates of the rival theory endorse another set of standards. If one theory is supported by one set of standards, and the rival by another, and there is no higher set of standards, then there would appear to be no basis for a rational choice between such theories. In the end, a radical epistemological relativism may seem unavoidable, since without a fixed method to arbitrate between rival sets of methodological standards, rationality can at best depend on whatever sets of standards a scientist happens to employ.

Non-Algorithmic Rationality

Numerous authors have reacted to the work of the historical school by arguing that the thesis of methodological variation leads to relativism and irrationalism.[1] Recently, however, some authors have suggested that, rather than leading to relativism and irrationalism, the work of the historical school leads to a new conception of scientific rationality.[2] I wish to suggest that such a new conception of scientific rationality is already to hand, and is available within the work of the historical school. To this end, I will present four characteristic theses of the historical school, which, taken together, yield a new model of scientific rationality.

The fundamental tenet on which this new model of scientific rationality rests is one of the leading themes of historical philosophy of science. It might even be taken as the thesis which unites it into a philosophical school. This is the thesis that rational choice between conflicting scientific theories cannot in general be governed by an algorithm of theory-choice. That is to say, the evaluative rules and criteria which make up the methodology of science cannot be fashioned into a single, universally acceptable, deterministic procedure, capable of being employed in a mechanical way to yield a unique choice between alternative scientific theories.

This thesis is succinctly expressed in the following quote from Kuhn's Postscript to the second edition of *The Structure of Scientific Revolutions*:

> There is no neutral algorithm for theory-choice, no systematic decision procedure which, properly applied, must lead each individual in the group to the same decision. (1970, p. 200)

This claim of Kuhn's is a negative existential claim, and as such its truth is unable to be conclusively established. Despite the inconclusiveness of this claim, however, I suggest that it should be interpreted as a generalization based on a study of past science, which has a great deal of historical plausibility. In particular, evidence from the history of science, both of past methodological change and of repeated and prolonged disagreement throughout the history of science, indicates that no algorithm for theory-choice has yet to be discovered. Moreover, in view of the complexity and variability of actual theory-choice situations, it is a highly plausible conjecture that no such algorithm is ever likely to be found.

It is important to note that Kuhn actually says that there is no *neutral* algorithm for theory-choice. Presumably, the reason for this qualification is that, while it may in fact be possible to formulate an algorithm capable of uniquely determining theory-choice, all such algorithms beg the question in favour of a particular theory or methodological criterion. It will simplify matters, however, to formulate this first thesis as follows:

T1: There is no algorithm of scientific theory-choice.

Before proceeding to the second thesis, two comments are in order about T1. First, the denial of an algorithm of theory-choice should not be understood as a denial of the existence of a scientific method or of a set of methodological criteria. What T1 denies is that there is any *universal* method or set of criteria which is capable of mechanically deciding between alternative theories. Second, T1 should also not be taken to deny that there are algorithmic rules which occur in science: for even if there is no single,

universal algorithm of theory-choice, there may still be individual rules which function as algorithms.

Methodological Pluralism

The second thesis which I propose is also principally due to Kuhn, though it is found in other authors as well.[3] According to this thesis, there is, instead of a single scientific method, an array of evaluative criteria to which scientists may appeal in choosing between theories. Kuhn lists as examples of such criteria, predictive accuracy, consistency, scope, simplicity and fertility (1977, pp. 321-322), and he comments that

> [such] criteria of choice ... function not as rules, which determine choice, but as values, which influence it. (1977, p. 331)

I will not follow Kuhn in using the term 'value' to refer to methodological criteria, though I do think Kuhn is right that there is a significant difference between those criteria which dictate the outcome of a decision and those which merely serve as a guide to choice.

In light of Kuhn's remarks, I propose the following statement of methodological pluralism as the second thesis of the present model:

> *T2:* In choosing between scientific theories, scientists draw upon an array of evaluative criteria, which guide or influence rather than determine their choice of theory.

Whereas T1 is a negative thesis which denies an algorithm of theory-choice, T2 is a positive thesis which asserts the pluralistic nature of scientific methodology. The positive thesis in T2 is, however, complementary to the denial in T1 of a universally applicable algorithm of theory-choice. For, while

there may be no algorithm of theory-choice, there may nevertheless be a range of evaluative criteria, which scientists employ in deciding which theory to accept.

According to the pluralistic model of scientific methodology, scientists have at their disposal a range of criteria of theory-appraisal, which jointly constitute the methodology of science. Since scientists may modify and replace methodological criteria in the advance of science, there may be variation in the set of criteria employed during the history of science. Similarly, since different branches of science may develop in different ways, there may also be methodological variation across the sciences. While it is not possible at this point to provide a complete taxonomy of such methodological criteria, the plausibility of the pluralist approach requires that at least a preliminary indication of such a taxonomy be given.

Evaluative criteria, of the sort discussed by methodological pluralists, range from general criteria and principles of theory-appraisal to specific rules of experimental procedure. Examples of the former might include the criteria mentioned by Kuhn (e.g., simplicity, coherence, accuracy), as well as Popper's dictum that scientists should maximize the falsifiability of theories by avoiding *ad hoc* hypotheses. As examples of the latter, one might think of instructions for proper use of instrumentation, procedures to insure purity of samples or accuracy of measurement, and so forth. Located somewhere between the extremes of general criteria and rules of laboratory practice, one would find criteria of explanatory adequacy, such as being a well-tested hypothesis, or being appropriately logically related to the phenomena to be

explained, as well as norms of proper test-procedure, such as repeatability and the use of experimental controls or blinds.

Two Corollaries of Pluralism

Turning to the two remaining elements of the model, both the third and fourth theses are, in effect, corollaries of T2. The third thesis stems from the observation that there may be conflict between the various evaluative criteria.[4] For, while it is in principle possible for a single theory to maximally satisfy all criteria, in practice conflicting theories might each satisfy different criteria better than their rivals. Hence,

> *T3:* The evaluative criteria employed in scientific theory-choice may conflict in application to alternative theories.

Where such a conflict between methodological criteria occurs, the set of such criteria cannot itself uniquely dictate the outcome for choice of theory. As such, the potential for conflict between criteria further exemplifies the claim in T1 that there is no algorithm of theory-choice.

The fourth thesis derives from Feyerabend's claim that all methodological rules have limitations, and are therefore defeasible. Such defeasibility of methodological criteria is the main lesson to be learned from Feyerabend's critique of scientific method. Feyerabend is, of course, famous for having claimed in *Against Method* that, as far as the methodology of science is concerned, "anything goes" (1975, p. 28). Yet it is often overlooked that the (non-rhetorical) force of this claim is not to deny that there are normative rules to which the practice of science conforms. Rather, it is to

deny that there are any inviolable rules of scientific methodology. As such, the point of the claim that "anything goes" is merely to jokingly concede, in case of insistence on a universal formulation of method, that, even in the absence of inviolable rules, there remains a universally applicable methodological rule -- namely, that expressed by the statement "anything goes".

Thus, in light of Feyerabend's critique of universal method, I propose the following thesis:

> *T4:* No evaluative criterion employed in scientific theory-choice is inviolable in all circumstances.

That thesis T4 is also a corollary of T2 can be seen from the fact that conflict between methodological criteria may make it necessary to decide between such criteria in order to choose between theories. But if a decision must be made on which of conflicting criteria to adhere to, then it follows that it must be possible to violate or override some criteria in favour of others.

It may at first appear that T4 is too strong. For to say that no criterion is inviolable appears to suggest that one need not follow any rule of scientific methodology in rational choice of theory: one need not do so because no rule is binding. There are at least three points to be made in reply to this objection. First, it should be noted that it does not follow from the violability of one criterion that all criteria are violable at the same time: that there are circumstances in which a rule may be broken does not entail that no rules need be followed. Second, it is consistent with asserting the violability of rules to also assert that at least minimal adherence is required to the set of

rules to which the violated rule belongs. While no single criterion need be followed, it would not be rational to accept a theory which violates all methodological criteria. Third, to assert the violability of criteria is not to assert the rationality of indiscriminate flouting of criteria. For while all criteria may be violable in some circumstances, there need nevertheless be good reasons for doing so: the circumstances must warrant such violation.

Solving the Problem of Theory-Choice

Taken together, theses T1-T4 constitute a non-algorithmic, pluralistic model of scientific methodology. On such a model, there is no single, universally acceptable procedure of theory-appraisal capable of dictating unique choice of theory, and scientists may appeal to a variety of different criteria in defending their preferred theoretical alternative. Let us see how this model solves the problem of rational theory-choice.

According to the present model, scientists confronted with a choice between alternative theories may take into account a range of different methodological criteria. One scientist might choose to accept a given theory, say the theory of continental drift in the early 20th century, because it provides the best available explanation of a broad range of phenomena (e.g., species distribution, geological pattern-matching, paleoclimatological data), which the scientist regards as particularly important. A second scientist might dismiss the drift hypothesis as unacceptably *ad hoc* (e.g., due to the absence of a suitable drift mechanism) and excessively speculative. Such a scientist might favour instead the theory that the continents are permanent

fixtures on the Earth's surface, which sought to explain geological phenomena without postulating any processes other than those for which there is direct empirical evidence (e.g., sedimentation, earthquakes, erosion). Yet a third scientist might reject such permanentism as an inadequate account of mountain-formation, and adopt instead some version of the contracting-Earth hypothesis. Contractionism derived support from physicists' claims of a cooling Earth, explained the formation of mountains as the crumpling of the Earth's crust as it gradually shrunk, and accounted for species distribution by means of the existence of land-bridges between continents at earlier periods of the Earth's history.[5]

On a scenario such as this, opposing scientists adopt different geological theories on the basis of divergent assessments of the epistemic merits of the competing theories. In support of their divergent assessments, scientists appeal to a variety of evaluative criteria, such as explanatory scope, *ad hoc*ness, empirical verifiability, and support from a related discipline. In so doing, they are able to marshall supporting arguments on behalf of their favoured theories on the basis of a diverse range of methodological criteria. As a result, opposing scientists may have rational grounds for choice of theory, in spite of adopting rival theories.

There may, in other words, be *rational disagreement* between scientists who accept conflicting theories on the basis of different methodological considerations. This is precisely what one would expect on a non-algorithmic, pluralist conception of scientific reason. For, in the absence of a single methodological procedure able to uniquely dictate choice of theory,

there is scope for scientists to arrive at a variety of divergent appraisals of the comparative strengths and weaknesses of competing theories.

As for the issue of whether divergent choice of theory may be rational in an *objective* sense, there seems no reason to suppose that divergent choice of theory based on variant methodological criteria need be lacking in objectivity. For, on the assumption that the rival scientists' choices are indeed made on the basis of appropriate methodological criteria, such choices would appear to be based on good reasons of a perfectly objective kind.

Relativism

It will, no doubt, be objected that the solution I have sketched to the problem of theory-choice is no solution at all. One scientist rationally accepts one theory on the basis of one set of methodological criteria. Other scientists rationally accept competing theories on the basis of other methodological criteria. Such variation of rational belief with methodological criteria is nothing short of relativism.

I wish to conclude this paper by briefly indicating why this objection seems to me to be incorrect. In the first place, it is a mistake to suppose that admitting that there may be divergent methodological grounds for conflicting choice of theory commits one to epistemic relativism. To be sure, the present conception of scientific reason contains a large measure of epistemic tolerance. But to tolerate divergence of rational belief is rather different from rendering such belief relative to operative standards. Such tolerance, moreover, is a necessity forced on us by even the most casual acquaintance

with the history of science, or indeed, with intellectual life generally. The phenomenon of rational disagreement is a robust phenomenon which must be taken into account by any theory of scientific rationality.

Why, then, does the idea of rational disagreement on the basis of alternative methodological criteria so readily elicit the charge of relativism? I suggest that the answer lies with an illegitimate assumption concerning the relation between methodological criteria and rational justification. For consider what would have to be the case in order for it to follow from variation in the criteria appealed to by scientists that rational theory-choice is relative to such criteria. In order for such choice to be relative to variant criteria, it would have to be the case that conformity with such criteria *suffices* for rational theory-choice. In other words, mere accordance with operative methodological criteria would be all that is needed for acceptance of a theory to be rationally justified.

Such an assumption is untenable, however, as may be seen by reflection on the fact that not all criteria which might be employed in theory-appraisal are able to provide genuine epistemic support. Appeal might, for example, be made to a purely aesthetic factor which has no bearing on the likely truth of the theory. Alternatively, use might be made of methodological criteria, which have been discredited, or which have been found wanting, as, for example, single-blind drug tests have been found wanting in light of the placebo effect.[6] Cases such as these reveal that there may be deficiencies in the evaluative criteria employed by scientists, which either preclude their having probative force, or which reduce such force as they might have.

The point I am making is based on a distinction between normative and descriptive issues at the level of methodological criteria. As far as the description of actual scientific practice is concerned, scientists may in fact employ a whole range of different criteria in their appraisal of alternative theories. Yet, as for the normative dimension of such practice, the criteria which scientists actually employ need not necessarily convey epistemic support. Since criteria may be deficient, or otherwise lacking in probative force, a scientist's belief may conform with a criterion without thereby being rationally justified.

Since the charge of relativism against the present model assumes that conformity with operative criteria insures rationality, it disregards the distinction between operative criteria and criteria which convey epistemic support. Yet if we insist on distinguishing descriptive from normative issues at the level of criteria, we may assert that scientists are guided by a plurality of divergent criteria without thereby licensing the inference to epistemic relativism. In particular, the model of rationality proposed here is fully consistent with enforcing a sharp distinction between criteria which provide epistemic support and criteria which provide no such support. The present model fails, therefore, to make theory-choice relative to operative criteria.

It might, finally, seem somewhat paradoxical to suppose that there may be disagreement between scientists whose beliefs are based on objective grounds. For it may seem to be part of the very concept of objectivity that it is bound up with consensus and convergence of belief. One might think, for example, that if there are objective grounds for a belief, then, necessarily, if

anyone is presented with such grounds, they should accept the belief. Given such a connection between objectivity and convergent belief, how can there be objectivity if scientists disagree?

The present model of rationality requires that objectivity be conceived as separate from consensus. For if there may be rational divergence between scientists whose beliefs are objectively grounded, then objectivity evidently cannot be bound up with agreement. The seeming paradox of this idea may be alleviated, however, if the locus of objectivity is situated in the criteria of evaluation themselves, rather than in the formation of consensus. For a scientist's acceptance of a theory may be objectively grounded if it is based on appeal to methodological criteria capable of yielding genuine epistemic support. Such criteria provide objective grounds for theory-acceptance since they provide epistemic support for the theory which is independent of irreducibly subjective factors, such as personal taste, whim or prejudice. Being objective, at least where rational belief is concerned, reduces to being epistemically well-founded.[7]

Endnotes

1. Cf. Gower (1988), Lakatos (1970), Popper, (1970), Siegel (1987), Worrall (1988).

2. Bernstein (1988), Brown (1988), Putnam (1981).

3. Cf. Chalmers (1990), Ellis (1990), Feyerabend (1975), Laudan (1984), Newton-Smith (1981).

4. For example, Kuhn (1977, pp. 323-324) argues that, while consistency with current theory favoured the geocentric system, considerations of simplicity tended to favour Copernicus.

5. I have drawn the general outlines of the views of the imaginary Earth scientists described here from Homer LeGrand's discussion of the situation in geology in the early twentieth century in his (1988).

6. Cf. Laudan (1984, pp. 38-39).

7. This paper was written while I held a Visiting Fellowship at the Center for Philosophy of Science, the University of Pittsburgh. I wish to thank the Center for its support of my research.

References

Bernstein, Richard (1988), *Beyond Objectivism and Relativism*, University of Pennsylvania Press, Philadelphia

Brown, Harold I. (1988), *Rationality*, Routledge, London

Chalmers, Alan (1990), *Science and its Fabrication*, University of Queensland Press, St. Lucia

Ellis, Brian (1990), *Truth and Objectivity*, Blackwell, Oxford

Feyerabend, Paul (1975), *Against Method*, New Left Books, London

Gower (1988), 'Chalmers on Method', *British Journal for the Philosophy of Science*, 39, pp. 59-65

Kuhn, Thomas S. (1970), *The Structure of Scientific Revolutions*, 2nd ed., University of Chicago Press, Chicago

Kuhn, Thomas S. (1977), *The Essential Tension*, University of Chicago Press, Chicago

Lakatos, Imre (1970), 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. E. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge

Laudan, Larry (1984), *Science and Values*, University of California Press, Berkeley

LeGrand, Homer (1988), *Drifting Continents and Shifting Theories*, Cambridge University Press, Cambridge

Newton-Smith, William H. (1981), *The Rationality of Science*, Routledge and Kegan Paul, London

Popper, Karl (1970), 'Normal Science and its Dangers', in I. Lakatos and A. E. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge

Putnam, Hilary (1981), *Reason, Truth and History*, Cambridge University Press, Cambridge

Siegel, Harvey (1987), *Relativism Refuted*, Reidel, Dordrecht

Worrall, John (1988), 'The Value of a Fixed Methodology', *British Journal for the Philosophy of Science*, 40, pp. 376-388