If this is correct, however, then my attempt briefly to sketch a certain line of historical interaction between philosophy and the sciences has resulted in a rather surprising dialectical twist. For taking the interaction between the history of science and the history of philosophy seriously has led us to a point where it now appears that the currently popular diagnosis of the failure of logical positivism (a diagnosis due largely to the work of Kuhn and his followers) is fundamentally misleading. Indeed, it now appears that the underlying philosophical motivations of the logical positivists cannot happily be described as either naively empiricist, naively formalist, or naively ahistorical. Their empiricism was qualified by, and, I believe, entirely subordinated to, an essentially Kantian preoccupation with the a priori framework within which alone empirical claims have a definite meaning in the first place. Their formalism rested on the idea, which itself evolved naturally from the important developments taking place in the formal sciences themselves, that this a priori framework for empirical knowledge must be specified within the radically new conception of formal logic due to Frege and Russell. Finally, although the logical positivists' preoccupation with the a priori did indeed thereby preclude them from using the history of science as a philosophical tool, this did not prevent them from recognizing the profound philosophical significance of conceptual revolutions in science. On the contrary, their effort to articulate a coherent conception of the *relativized* a priori must, I think, count as the most rigorous attempt we have yet seen philosophically to come to terms with precisely such conceptual revolutions. Of course, as we have also seen, this heroic attempt of the logical positivists was in the end a failure. Yet I do not myself think that we will ever progress beyond this point until we possess a fuller appreciation of the historical evolution of our own philosophical predicament. And this means, as I have tried to emphasize throughout, that we must attend more closely to the history of science, the history of philosophy, *and* to the essential interaction between them.

Rationality and Paradigm Change in Science

Ernan McMullin

As we look back at the first responses of philosophers of science to Thomas Kuhn's classic *The Structure of Scientific Revolutions,* we are struck by their near unanimity toward the challenge that the book posed to the rationality of science. Kuhn's account of the paradigm changes that for him constituted scientific revolutions was taken by many to undermine the rationality of the scientific process itself. The metaphors of conversion and gestalt switch, the insistence that defenders of rival paradigms must inevitably fail to make contact with each other's viewpoints, struck those philosophical readers whose expectations were formed by later logical empiricism as a deliberate rejection of the basic requirements of effective reason giving in the natural sciences.

Kuhn responded to this reading of *SSR* in a lengthy Postscript to the second edition of his book in 1970 and in the reflective essay "Objectivity, Value Judgement, and Theory Choice" in 1977.[1] He labored to show that the implications of his new account of scientific change for the *rationality* of that change were far less radical than his critics were taking them to be. But his disavowals were not, in the main, taken as seriously as he had hoped they would be; the echoes of the rhetoric of *SSR* still lingered in people's minds. It seems worth returning to this ground, familiar though it may seem, in order to assess just what Kuhn *did* have to say about how paradigm change comes about in science. We will see that the radical thrust of his account of science was indeed not directed so much against the rationality of theory choice as against the epistemic, or truthlike, character of the theories so chosen.

1 Good Reasons for Paradigm Change

The theme that recurs in Kuhn's discussions of paradigm change is a two-sided one. On one hand, he wanted to emphasize the fundamental role played by "good reasons" in motivating theory change in science. Notable among these is the perception of anomaly, the growing awareness that something is wrong, which makes it possible for alternatives to be seriously viewed *as* alternatives. On the other hand, these reasons are never coercive in their own right in forcing change; the reasons in favor of a new paradigm cannot *compel* assent. There is no precise point at which resistance to the change of paradigm becomes illogical.[2] Proponents of the new paradigm and defenders of the old one may each be able to lay claim to be acting "rationally"; the fact that neither side can persuade the other does not undermine the claim each can make to have good reasons for what they assert. "The point I have been trying to make," Kuhn says in the Postscript to *SSR,* "is a simple one, long familiar in philosophy of science. Debates over theory-choice cannot be cast in a form that fully resembles logical or mathematical proof. . . . Nothing about that relatively familiar thesis implies either that there are no good reasons for being persuaded or that those reasons are not ultimately decisive for the group. Nor does it even imply that the reasons for choice are different from those usually listed by philosophers of science: accuracy, simplicity, fruitfulness, and the like. What it should suggest, however, is that such reasons function as values and that they can thus be differently applied, individually and collectively, by men who concur in honoring them."[3]

It is with the implications of this thesis that I will be mainly concerned in this essay. The values a good theory is expected to embody enable comparisons to be made, even when the rival theories are incommensurable. Kuhn makes it clear that "incommensurable" for him does not imply "incomparable." *SSR,* he notes, "includes many explicit examples of comparisons between successive theories. I have never doubted either that they were possible or that they were essential at times of theory choice."[4] What he wanted to emphasize, he

says, is that "successive theories are incommensurable (which is not the same as incomparable) in the sense that the referents of some of the terms which occur in both are a function of the theory within which those terms appear," and hence that there is no neutral language available for purposes of comparison. Nonetheless, translation is in principle possible.[5] But to translate another's theory is still not to make it one's own. "For that one must go native, discover that one is thinking and working in, not simply translating out of, a language that was previously foreign."[6] And that transition cannot simply be willed, he maintained, however strong the reasons for it may be. This is what enabled him to maintain his most characteristic claim, even after the qualifiers he inserted in the Postscript: "The conversion experience that I have likened to a gestalt switch remains, therefore, at the heart of the revolutionary process. Good reasons for choice provide motives for conversion and a climate in which it is more likely to occur. Translation may, in addition, provide points of entry for the neural reprogramming that, however inscrutable at this time, must underlie conversion. But neither good reasons nor translation constitute conversion, and it is that process we must explicate in order to understand an essential sort of scientific change."[7]

How is the transition to be explicated? Kuhn has only some hints to offer: "With respect to divergences of this sort, no set of choice criteria yet proposed is of any use. One can explain, as the historian characteristically does, why particular men made particular choices at particular times. But for that purpose one must go beyond the list of shared criteria to characteristics of the individuals who make the choice. One must, that is, deal with characteristics which vary from one scientist to another without thereby in the least jeopardizing their adherence to the canons that make science scientific."[8]

And he mentions such characteristics as previous experience as a scientist, philosophical views, personality differences. In the years since *SSR* appeared, sociologists of science have made much of these factors, often in ways that Kuhn himself would disavow. It was his stress on the role of these factors, he later remarked, that led critics to dub his views "subjectivist." They forgot his stress on the "shared criteria" that guide (but do not dictate) theory choice.[9] I will take him at his word here, assuming that the rationality of theory choice in his account rests on the persistence of these criteria that enable theories to be compared and evaluated, relatively to one another, even when they are incommensurable.

2 How Deep Do Revolutions Go?

Here we immediately encounter a difficulty. *Do* these criteria persist? Can they bridge paradigm differences? How deep, in short, do revolutions go? There is an ambiguity in Kuhn's response to this question. In a celebrated paragraph in *SSR,* he describes paradigm change as follows: "Like the choice between competing political institutions, that between competing paradigms proves to be a choice between incompatible modes of community life. Because it has that character, the choice is not and cannot be determined merely by the valuative procedures characteristic of normal science, for these depend in part upon a particular paradigm, and that paradigm is at issue. When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular. Each group uses its own paradigm to argue in that paradigm's defense."[10]

Since the evaluative procedures depend on the paradigm, and the paradigm itself is in question, there can be no agreed-upon way to adjudicate the choice between rival paradigms. Though he goes on to say that the resulting circularity does not *necessarily* undercut the arguments used, he concludes that the status of such arguments can at best be only that of persuasion. They "cannot be made logically or even probabilistically compelling for those who refuse to step into the circle. The premises and values shared by the two parties to a debate over paradigms are not sufficiently extensive for that."[11]

What prevents the rival parties from agreeing as to which paradigm is the better, then, is in part the fact that the norms in terms of which this debate could be carried on are themselves part of the paradigm, so that there is no neutral methodological ground, or at least not enough to enable agreement to be reached. How important is this sort of "circularity" to Kuhn's account of the inability of either side in a paradigm debate to muster an entirely cogent argument in its own behalf? If a circularity in regard to evaluative procedures were to hold in general in such cases, then scientific revolutions *would* indeed seem to be the irrational, or at least minimally rational, affairs that Kuhn's critics take him to be saying they are. One way to find out is to direct attention to the examples he gives of scientific revolutions and ask what paradigm change amounts to in each of these cases.

When the question is put in this way, it is clear that there is a striking difference in the depth of the different changes classified by Kuhn as "revolutions." At one end of the spectrum is the Copernican revolution, the charting of which led him to the writing of *SSR* in the first place. At the other end would be, for example, the discovery of X rays. Somewhere in the middle might come the discovery of the oxygen theory of combustion.[12] We have a choice in some cases, it would seem, between saying that only a small part of the paradigm changed and saying that an entire paradigm changed but that the "paradigm" in this case comprised only a fraction of the beliefs, procedures, and so forth, of the scientists involved.

Take the case of X rays. Kuhn insists that their discovery did accomplish a revolution in his sense. Yet he recognizes that at first sight this episode scarcely seems to qualify. After all, no fundamental change of theory occurred. No troublesome anomalies were noted in advance. There was no prior crisis to signal that a revolution might be at hand. Why then, he asks, can we not regard the discovery of X rays as a simple extension of the range of electromagnetic phenomena? Because, he responds, it "violated deeply entrenched expectations . . . implicit in the design and interpretation of established laboratory procedures."[13] The use of a particular apparatus "carries with it the assumption that only certain sorts of circumstances will arise." Roentgen's discovery "denied previously paradigmatic types of instrumentation their right to that title." That was sufficient, in his view, to constitute it a "revolution" in the sense in which he is proposing to use that term.

I will call this a shallow revolution because so much was left untouched by it. Electromagnetic theory was not replaced or even altered in any significant way. There were no challenges to accepted ways of assessing theory or to what counts as proper explanation. The textbooks, the sets of approved problem solutions, did not change much. What changed were the experimental procedures used in working with cathode-ray equipment and the expected outcomes of such work. And, of course, there were some important long-range implications for theory (as we now know). Such "revolutions" ought, it would seem, to be fairly frequent. Much would depend on how literally one should take the criteria Kuhn specifies as being the symptoms of impending revolution: previous awareness of anomaly and a resistance to a threatened change in procedures or categories.[14]

We are much more likely to think in terms of "revolution" in cases where one large-scale theory replaces another. Kuhn's favorite example is the replacement of phlogiston theory by the oxygen theory of combustion.[15] It meant a reformulation of the entire field of chemistry, a new conceptual framework, a new set of problems. Another example he gives of this sort of intermediate revolution, as we might call it, is the discovery of the Leyden jar and the resulting emergence of "the first full paradigm for electricity."[16] Prior to this discovery, Kuhn remarks, no single paradigm governed electrical research. A number of partial theories were applied, none of them entirely successful. The new conceptual framework enabled normal science to get under way, even though one-fluid and two-fluid theories were still in competition.

These changes involved the formulation of a new and more comprehensive theory. But they left more or less unchanged the epistemic principles governing the paradigm debate itself. Both sides would have agreed as to what counts as evidence, as to how claims should be tested. Or more accurately, to the extent that the scientists involved would have disagreed on these issues, their disagreements would not have been paradigm-dependent to any significant extent. So far as we can tell, Priestley and Lavoisier applied the same sorts of criteria to the assessment of theory, though they might not have attached the same weight to each criterion.

In Kuhn's favorite example of a scientific revolution, the Copernican one, this was, of course, not the case. This was a revolution of a much more fundamental sort because it involved a change in what counted as a good theory, in the procedures of justification themselves. It was not abrupt; indeed, it took a century and a half, from Copernicus's *De revolutionibus* to Newton's *Principia,* to consummate. And what made it revolutionary was not just the separation of Newtonian cosmology or Newtonian mechanics from their Aristotelian counterparts but the gradual transformation in the very idea of what constitutes valid evidence for a claim about the natural world, as well as in people's beliefs about how that world is ordered at the most fundamental level.[17]

It can thus be called a *deep* revolution, by contrast with the others described above. The Aristotelians and the Galileans totally disagreed as to how agreement itself should be brought about. So did the Cartesians and the Newtonians. The Galileans made use of idealization, of measurement, of mathematics, in ways the Aristotelians believed were illegitimate. The Newtonians allowed a form of explanation that the Cartesians were quite sure was improper. The shift in paradigm here meant a radical shift in the methodology of paradigm debate itself. Paradigm replacement means something much more thoroughgoing in such a case.

Have there been other deep revolutions in the more recent history of natural science? Newton's success means the success of a methodology which is still roughly the methodology of natural science today. Perhaps only one deep revolution was needed to get us to what Kuhn calls "mature" science. The two major revolutions in the physics of our own century did not run quite so deep. But they *did* involve principles of natural order, that is, shared assumptions as to what count as acceptable ways of articulating physical process at its most basic level. In the quantum revolution, what separated Bohr and Einstein was not just a difference in theoretical perspective but a disagreement as to what counted as good science and why. Quantum theory, in its Copenhagen interpretation, came much closer to a deep paradigm replacement than it would have done in Einstein's way of taking it.

In the Postscript to *SSR,* Kuhn addressed the ambiguity of the notion of paradigm and proposed a new label. A disciplinary matrix is the answer to the question, "What does [a community of specialists] share that accounts for the relative fullness of their professional communication and the relative unanimity of their professional judgments?"[18] Some of its principal components, he says, are symbolic generalizations, models of the underlying ontology of the field under investigation, concrete problem solutions, and the values governing theory appraisal.

It is clear, then, that for there to be a revolution in Kuhn's sense of the term this last component does not have to be at issue. Only in a deep revolution does one side challenge the other in regard to the appropriate methodology of theory assessment. When X rays were discovered, there was no dispute as to how their reality should be tested. When a Kuhnian revolution takes place, it is evidently not necessary that the entire paradigm should change. Only a part of the disciplinary matrix need be affected for there to be a sufficient change in worldview to qualify as "revolutionary." What 'revolutionary' means in practice is a change that falls outside the normal range of puzzle-solving techniques and whose resolution cannot, therefore, be brought about by the ordinary resources of the paradigm.

The implicit contrast is between puzzle solving, with its definitive ways of deciding whether a puzzle really *is* solved, and paradigm debate, where no such means of ready resolution exists. Whether so sharp a contrast is warranted by the actual practice of science may well be questioned. Decision between rival theories is an everyday affair in any active part of science. There may be an accepted general framework within which problems are formulated, but new data constantly pose challenges to older subtheories within that framework. This was the main issue dividing Kuhn and his Popperian critics in the late 1960s. It is clear in retrospect that there was merit on both sides of that dispute but that each was focusing on a particular aspect of scientific change to the exclusion of others.

The appraisal of rival theories within a paradigm is not a simple matter of puzzle solving. The history of high-energy physics over the past thirty years, for example, has seen one theory dispute after another. The notorious divisions at the moment among paleontologists about the causes of the Cretaceous extinction or between planetary physicists about the origin of the moon are only two of the more obvious reminders of the fact that deep-seated disagreement about the merits of alternative theories is a routine feature of science at its most "normal." As we have seen, Kuhn traced the roots of paradigm disagreement to two different sources: an "incommensurability" of a complex sort between two ways of looking at the world and a set of criteria for theory choice that function as values to be maximized rather than as an effective logic of decision. But this latter source of difference characterizes theory disputes generally and not just the more intractable ones that Kuhn terms paradigm disagreements. What we have here, I suspect, is a spectrum of different levels of intractability, not just a sharp dichotomy between revolutions and

puzzle solutions. Nevertheless, Kuhn's dichotomy, though rather idealized, did serve to bring out in a forceful and dramatic way how complex, and how far from a simple matter of demonstration, the choice between theoretical alternatives ordinarily is.

3 The Virtues of a Good Theory

What makes this choice a *rational* one for Kuhn, as we have seen, is the fact that scientists are guided by what they would regard as the virtues of a good theory. And there has been a certain constancy in that regard, according to him, across all but perhaps the deepest of revolutions: "I have implicitly assumed that, whatever their initial source, the criteria or values deployed in theory choice are fixed once and for all, unaffected by their transitions from one theory to another. Roughly speaking, but only roughly speaking, I take that to be the case. If the list of relevant values be kept short (I have mentioned five, not all independent) and if their specification is left vague, then such values as accuracy, scope, and fruitfulness are permanent attributes of science."[19]

This is a strong assertion indeed. Ironically, it is stronger than that now made by some of those who, like Laudan and Shapere, have chided Kuhn in the past for his subjectivism.[20] They argue that the values involved in theory choice are in no sense fixed; Shapere objects to any such claim as an objectionable form of essentialism. According to Laudan and Shapere, these values themselves change gradually as theories change or are replaced. They change for *reasons,* they insist, these reasons functioning as some sort of higher-level arbitration. But there is no limit in principle as to how *much* they might change over time. To put this in a more direct way, there is no constraint on how different the criteria of a good theory might be in the science of the far future from those we rely on today, unlikely though a radical shift might be.[21] In the original text of *SSR,* Kuhn proposed what sounds like a rather different view:

> When paradigms change, there are usually significant shifts in the criteria determining the legitimacy both of problems and of proposed solutions. . . . [This is] why the choice between competing paradigms regularly raises questions that cannot be resolved by the criteria of normal science. To the extent, as significant as it is incomplete, that two scientific schools disagree about

what is a problem and what a solution, they will inevitably talk through each other when debating the relative merits of their respective paradigms. In the partially circular arguments that regularly result, each paradigm will be shown to satisfy more or less the criteria that it dictates for itself and to fall short of a few of those dictated by its opponent.[22]

The criteria governing theory choice are described here as strongly paradigm-dependent and thus as suffering "significant shifts" from one paradigm to the next. The resulting partial circularity in paradigm assessment leads rival scientists to "talk through each other." This was the theme, of course, that Paul Feyerabend picked up on. One can see how severely it limits the notion that there are "good reasons" for paradigm change. Here, then, is a clear instance of how Kuhn's later construals soften the radical overtones of the earlier work.

Kuhn does not hesitate to speak of the values involved in theory appraisal as "permanent attributes of science." He allows that the manner in which these values are understood and the relative weights attached to them have changed in the past and may change again in the future. But he wants to emphasize that these changes at the metalevel tend to be slower and smaller in scale than the changes that can occur at the level of theory:

If such value changes had occurred as rapidly or been as complete as the theory changes to which they related, then theory choice would be value choice, and neither could provide justification for the other. But, historically, value change is ordinarily a belated and largely unconscious concomitant of theory choice, and the former's magnitude is regularly smaller than the latter's. For the functions I have here ascribed to values, such relative stability provides a sufficient basis. The existence of a feedback loop through which theory change affects the values which led to that change does not make the decision process circular in any damaging sense.[23]

One would need, however, to know just how and why changes in theory bring about changes at the metalevel of theory assessment in order to judge how large these latter changes might become without undermining the claim that a rational choice is being made. Is the "relative stability" of the criteria governing theory choice a contingent historical finding, or is it a necessary feature of any activity claiming the title of science? There are suggestions of both views in the passage I have just quoted. Historically, these values have in fact been stable,

Kuhn remarks. But he adds that if they were not, if one had to choose the criteria of choice themselves in the act of choosing between theories, there would be no fulcrum. The process would lack justification; it would be circular in a way that would be damaging to its claim to qualify as science.

The presumption appears to be that *really* deep revolutions do not occur, that is, revolutions where there is *no* sharing of epistemic values between one paradigm and the other. Kuhn allows that large-scale theory change may involve smaller-scale changes in the values believed to be appropriate to theory appraisal. In such cases, adoption of the new paradigm carries with it adoption of a somewhat different "rationality" at the metalevel. The advantages of the new theory are so marked, in terms of a minimal level of shared values, that a shift in the values themselves is ultimately taken to be warranted. This, it can be argued, is what happened in the seventeenth century as the balance shifted between Aristotelians and Galileans. Galileo set out to undermine Aristotle's physics in its *own* terms first and then to present an alternative that, in terms of consistency, empirical adequacy, and future potential, could claim a definite advantage, even in terms of criteria the Aristotelian might be brought to admit. That, at any rate, would be the grounds, in Kuhn's perspective, for regarding the Scientific Revolution as a "rational" shift in the way in which natural science was carried on.

In a recent essay Kuhn argues that we learn to use the term 'science' in conjunction with a cluster of other terms like 'art', 'medicine', 'philosophy'. To know what science is, is to know how it relates to these other activities.[24] Identifying an activity as scientific is to single out "such dimensions as accuracy, beauty, predictive power, normativeness, generality, and so on. Though a given sample of activity can be referred to under many descriptions, only those cast in this vocabulary of disciplinary characteristics permit its identification as, say, science; for that vocabulary alone can locate the activity close to other scientific disciplines and at a distance from disciplines other than science. That position, in turn, is a necessary property of all referents of the modern term, 'science.'"[25]

He immediately qualifies this last very strong claim by noting that not every activity that qualifies as "scientific" need be predictive, not all need be experimental, and so forth. And there is no sharp line of demarcation between science and nonscience. Nonetheless, there is a well-defined cluster of values whose pursuit marks off scientific from other activities in a relatively unambiguous way and that gives the term 'science' the position it occupies in the "semantic field." This marking off is not a mere matter of convention. The taxonomy of disciplines has developed in an empirical way; a real learning has taken place. If someone were to deny the rationality of learning from experience, we would not know what he or she is trying to say. One cannot, he maintains, further *justify* the norms for rational theory choice. He cites C. G. Hempel to the effect that this inability is a testimony to our continuing failure to solve the classical problem of induction.[26]

Kuhn rests his case, then, both for the rationality of science and for its distinctiveness as a human activity mainly on the values governing theory choice in science. But he does not chronicle their history, disentangle them from one another except in a cursory way, or inquire in any detail into how and why they have changed in the ways they have. Many of these variations, he remarks, "have been associated with particular changes in scientific theory. Though the experience of scientists provides no philosophical justification for the values they deploy (such justification would solve the problem of induction), those values are in part learned from that experience, and they evolve with it."[27]

But what justification other than the experience of scientists is *needed* to justify the values they deploy? Kuhn has, I suspect, altogether too lofty a view of what "philosophical" justification might amount to. And he has too readily allowed himself to be intimidated by that most dire of philosophers' threats: "That *can't* be right: if it were, it would solve the problem of induction." My own guess is that attention to the role of values in theory appraisal might well dissolve the problem Hume bequeathed us about the grounds for inductive inference. But whether that be true or not, the criteria employed by scientists in theory evaluation enjoy whatever sanction is appropriate to something learned in, and tested by, experience.

4 How Might Epistemic Values Be Validated?

Suppose a scientist were to doubt whether a particular value, say simplicity, is really a desideratum in a practical situation of theory

choice facing him or her. The rationality of the choice depends, presumably, on what sort of answer can be given to this kind of question. Two different sorts of answers suggest themselves. One is to look at the track record and decide how good a guide simplicity has proved to be in the past. (There are obvious problems about how the criterion itself is to be understood, but I will bracket these for the moment.) A quite different sort of response would be that simplicity is clearly a desideratum of theory because ____, where we fill the blank with a reason why on the face of it, a simple theory is more likely to be a good theory (if indeed one *can* find a convincing reason). Both of these responses would, of course, need further clarification before they could begin to carry any conviction.

First, what does it mean to ask how good a guide simplicity has been in the past? Guide to what? Some kind of ordering of means and ends is clearly needed here. Some of the values we have been talking about seem to function as goals of the scientific enterprise itself: predictive accuracy (empirical adequacy) and explanatory power are the most obvious candidates. One can trace each of these goals back a very long way in human history. In some sense, they may be as old as humanity itself. The story of how they developed in the ancient world, how the skills of prediction came to be prized in many domains, how explanatory accounts of natural process came to be constructed, is a familiar one. Less familiar is the realization that these goals were not linked together in any organic way at the beginning. Indeed, they were long considered antithetical in the domain of astronomy, the most highly developed part of the knowledge of nature in early times. One of the consequences, perhaps the most important consequence, of the Copernican revolution was to show that they *are* compatible, that they can be successfully blended. This was an empirical discovery about the sort of universe we live in. It was something we *learned* and that now we *know*.

Each of these goals has come to be considered valuable in its own right, an end in itself.[28] An activity that gives us accurate knowledge of the world we live in and consequently power over its processes can come to seem worthwhile for all sorts of reasons. An activity that allows us to understand natural process, that allows our imaginations to reach out to realms inaccessible to our senses, holds immediate attraction. What it is to understand will, of course, shift as the principles of natural order themselves shift. So this goal of explaining lacks the definiteness of the goal of predicting; as theory changes, so will the contours of what counts as explaining.

Much more would have to be said about all this, but I am going to press on to make my main point.[29] Other epistemic values serve as *means* to these ends; they help to identify theories more likely to predict well or to explain. Some of these are quite general and would apply to any epistemic activity. Logical consistency (absence of contradiction) and compatibility with other accepted knowledge claims would be among these. They are obviously not goals in themselves; they would not motivate us to carry on an activity in the first place. But we have found that these values are worth taking seriously as *means*. Or should I say, it has always been obvious that we must not neglect them, if it is knowledge we are seeking?

Other values are more specific to science, for example, fertility, unifying power, and coherence (i.e., absence of ad hoc features). Once again, these are clearly not primary goals. They are not so much deliberately aimed at as esteemed when present. And they are esteemed not in themselves but because they have proved to be the marks of a "good" theory, a theory that will serve well in prediction and explanation. A long story could be told about this, beginning with Kepler, Boyle, and Huygens and working through Herschel, Whewell, and a legion of others who have drawn attention to the significance of these three virtues.

Once again, the story is an ambiguous one: it can be told in two quite different ways. According to one way of telling it, these values can be shown to have played a positive historical role in theory choice; we have gradually learned to trust them as clues. According to the other, a series of acute thinkers (some of the most prominent of them listed above) have realized that these values *ought* to serve as indicators of a good theory. These are what one would *expect* a priori from a theory that purported to predict accurately and explain correctly. When Kepler and Boyle drew attention to the importance of such criteria, it was not to point to their efficacy in the earlier history of natural philosophy but to recommend them on general epistemic grounds.[30]

The question of how to validate the values that customarily guide scientific theory choice can now be addressed more directly. The goals of predictive accuracy (empirical adequacy) and explanatory power serve to define the activity of science itself, in part at least. If, as Kuhn notes, one relinquishes the goal of producing an accurate account of natural regularity, the activity one is engaged in may be worthwhile, but it is not science.[31] The notion of epistemic justification does not directly apply to the goals themselves. One might ask, of course, whether the pursuit of these goals is justifiable on *moral* grounds. Or one might ask, as a means of determining whether effort expended on them is worthwhile, whether the goals are in fact attainable. We have learned that in general they *are* attainable. This is something one could not have known a priori. And we have learned much about the *methods* that have to be followed for theory construction to get under way, methods of experiment, of conceptual idealization, of mathematical formulation, and the rest. All of this had to be *learned*, and no doubt there is still much to discover in this regard.

The other values, being instrumental, are justified when it is shown that they serve as means to the ends defined by the primary goals. And this, as we have seen, can be done in two ways: by an appeal to what we have learned from the actual practice of science or by an analysis in epistemological terms of the aims of theory and what, in consequence, the marks of a good theory should be. Ideally, both ways need to be followed, each serving as check for the other. The appeal to historical practice works not so much as a testimony to what values have actually guided scientists in their theory choices but as a finding that reliance on certain values has *in fact* served the primary goals of science. Might it cease to?

This is the Humean echo that seems to worry Kuhn so much. One might respond, as he does, that learning from experience is part of what it is to be rational. We cannot *demonstrate* that experience will continue to serve as a reliable guide. But demonstration is not what is called for. Kuhn has done more than anyone else, perhaps, to show that rational theory choice does not require the cogency of demonstration. We know that the predictive powers of natural science have enormously increased, and we know something of the theory characteristics that have served to promote this expansion. No future development could, so far as I can see, lead us to deny these knowledge claims, which rest not just on a perception of past regularities but on an understanding, partial at least, of why these regularities took the course they did. We can, and almost surely will, learn more about what to look for in a good theory. But no further evidence seems to be needed to show that coherence in a theory is a value to be sought, so that, other things being equal, a more coherent theory is to be preferred to a less coherent one.

5 Rationality without Realism?

Over the years since *SSR* appeared, Kuhn has, as we have seen, become more and more explicit about the basic rationality that underlies theory choice in science. It is a complex rationality with many components, allowing much latitude for difference among the defenders of different theories. But it has remained relatively invariant since the deep revolution that brought it into clear focus in the seventeenth century. One might almost speak of a *convergence* here. Kuhn clearly believes that scientists have a pretty good grip on the values that *ought* to guide the appraisal of rival theories, and that this grip has improved as it has been tested against a wider and wider variety of circumstances.

But he has not softened his stance in regard to the truth character of theories in the least. In a well-known passage in the Postscript, he insists that the only sort of progress that science exhibits is in puzzle solving: later theories solve more puzzles than earlier ones, or (to put this in a different idiom) they predict better. But there is, he insists, "no coherent direction of ontological development"; there is no reason to think that successive theories approximate more and more closely to the truth.[32] "The notion of a match between the ontology of a theory and its 'real' counterpart in nature now seems to me illusive in principle."[33] Kuhn thus rejects in a most emphatic way the traditional realist view that the explanatory success of a theory gives reason to believe that entities like those postulated by the theory exist, i.e., that the theory is at least approximately true.

He does not argue for this position in SSR, aside from a remark about Einstein's physics being closer in some respects to Aristotle's than to Newton's. But it is clear what the grounds for it are in his mind: the incommensurability of successive paradigms implies a discontinuity between their ontologies. By separating the issues of comparability and commensurability, he believes he can retain a more or less traditional view in regard to the former while adopting an instrumentalist one in regard to the latter. The radical challenge of *SSR* is directed not at rationality but at realism. The implications of the familiar Kuhnian themes of holism and paradigm replacement are now seen to be more significant for the debate about realism than for the issue of scientific rationality, on which they had so great an initial impact.

Kuhn's influence on the burgeoning antirealism of the last two decades can scarcely be overestimated. His views on theory change, on problems about the continuity of reference, are reflected in the work of such notable critics of realism as Arthur Fine, Bas van Fraassen, and especially Larry Laudan.[34] Kuhn's own emphasis on science as a puzzle-solving enterprise would lead one to interpret him in an instrumentalist manner. At this point I am obviously not going to open a full-scale debate on realism versus instrumentalism.[35] But I would like to pull out one thread from that notorious tangle. Kuhn's way of securing scientific rationality by focusing on the values proper to theory choice might well have led him (I argue) to a more sympathetic appreciation of realism. I am not saying that rationality and realism are all of a piece, that to defend one is to commit oneself to the other. Most of the current critics of realism would be emphatic in their defense of the overall rationality of scientific change. But a closer study of the values to which Kuhn so effectively drew attention should, to my mind, raise a serious question about the adequacy of an instrumentalist construal of the puzzle-solving metaphor. If such a construal is adopted, it is hard to make sense of those many episodes in the history of science where values other than mere predictive accuracy played a decisive role in the choice between theories.

To show this, I will focus on a case history from Kuhn's own earlier work, *The Copernican Revolution.* At issue are the relative merits of the Ptolemaic and the Copernican systems prior to Galileo's work. Kuhn points out that there was little to choose between the two on the score of predictive accuracy. "Judged on purely practical grounds," he concludes, "the Copernican system was a failure; it was neither more accurate nor significantly simpler than its Ptolemaic predecessors."[36] Yet it persuaded some of the best astronomers of the time. And it was they who ultimately produced the "simple and accurate" account that carried the day. How *did* it persuade them?

In Kuhn's view, "The real appeal of sun-centered astronomy was aesthetic rather than pragmatic. To astronomers the initial choice between Copernicus' system and Ptolemy's could only be a matter of taste, and matters of taste are the most difficult of all to define or debate."[37]

But such matters cannot be regarded as unimportant, he goes on, as the success of the Copernican Revolution itself testifies. Whatever it was that persuaded so many of those most skilled in astronomy to make what we would now regard as the right step obviously must be looked at with care. Those who were equipped "to discern geometric harmonies" obviously found "a new neatness and harmony" in the heliocentric system. What Copernicus offered was "a new and aesthetic harmony" that somehow carried conviction in the right quarters.

But now let us see how Copernicus's own argument went, in the crucial chapter 10 of book 1 of *De revolutionibus*. He points to two different sorts of clues. First, the heliocentric model allows one to specify the order of the planets outward from the central body in an unequivocal way, which Ptolemy's model could not do. Furthermore, the Copernican model has the planetary periods increase as one moves outward from the sun, just as one would expect. What Copernicus claims to discover in the new way of ordering the planets is a "clear bond of harmony," "an admirable symmetry." But why should this carry conviction, especially since (as Kuhn emphasizes) Copernicus in the end had to retain an inelegant and far from harmonious-seeming tangle of epicycles?

He had stronger arguments. The heliocentric model could *explain*, that is, provide the *cause* of, a whole series of features of the planetary motions that Ptolemy simply had to postulate as given, as inexplicable in their own right. For example, even in ancient times it had been suggested that Venus and Mercury appear to have the sun as their center of rotation, since, unlike the other planets, they accompany the sun in its motion across our sky. Or again, it had long been noted that the superior planets (Mars, Jupiter, Saturn) are at their brightest when in opposition (rising together in the evening or setting together in the morning). Assuming that brightness is a measure of relative distance, this is explained if we are viewing the planetary motions from a body that itself is orbiting the sun as center. This "proves,"

Copernicus somewhat optimistically concludes, that the center of motion of the superior planets is the same as that of the inferior planets, namely the sun.

Kuhn comments that it does "not actually prove a thing. The Ptolemaic system explains these phenomena as completely as the Copernican," although the latter can be said to be "more natural."[38] Here I must disagree. The Ptolemaic system does not *explain* the phenomena mentioned above at all. Ptolemy is forced to postulate that the center of the epicycle for both Venus and Mercury always lies on the line joining the earth and sun. Kuhn says that in this way Ptolemy "accounts for" this feature of their motions. But this is surely not *accounting for* in the sense of explaining. Kuhn evidently equates prediction and explanation in these passages, not an unusual assumption at the time his book was written.

But he allows that Copernicus gives a "far more natural" account than does Ptolemy. Why? And what does 'natural' mean in the lexicon of an instrumentalist? Ptolemy's restriction on the deferent radii swept out by Venus and Mercury "is an 'extra' device, an *ad hoc* addition,"[39] one that Copernicus can discard. Kuhn is surely on the right track here. But this is *not* an aesthetic argument, an appeal to taste. Copernicus himself makes the genre to which it belongs quite clear. He says that Copernicus is able to assign the *cause* of these features of the planetary motions, whereas Ptolemy is not. There is no reason in Ptolemy's system for them, other than the mere need to get the predictions right. They are, as Kuhn himself says, ad hoc.

Copernicus gives another set of arguments based on the retrograde motions. Their relative size and frequency from one planet to another and the lack of any such motions on the part of the sun and moon are exactly what one would be led to expect in a system where we are observing the motions from the third planet and the moon is not a true planet but a satellite of earth. Later, in the *Mysterium cosmographicum*, Kepler developed these arguments more fully and added some of his own, for example, the striking fact that in the Ptolemaic model, the period of rotation for each planet on either the deferent or the epicycle circle is *exactly* one year, something which seemed like an extraordinary piece of adjustment, especially since Ptolemy took the planets to be dynamically independent of one another. Kepler is clear that the issue here is one of causal explanation; one of the systems can provide such an explanation, the other cannot. He is also clear that the criterion of prediction alone will not be enough to decide in all cases between two rival accounts of the planetary motions and thus that a different genre of argument (he calls it "physical") is needed.[40] This he urged as a refutation of the instrumentalism of his opponent, Ursus.

The competition may have been neutral between Ptolemy and Copernicus where *prediction* of planetary motions was concerned, but the two systems were quite unequal as *explanation*. No better illustration could be found of the distinction between these two concepts, and of the consequent importance of criteria of theory appraisal other than that of predictive or descriptive accuracy. Copernicus's criterion of "naturalness," the elimination of ad hoc features, the virtue that might today be called coherence, is not aesthetic; it is epistemic. He is not just appealing to his reader's taste, or sense of elegance. He is not assuming that the simpler, the more beautiful, models are more likely to be true. He is saying that a theory that makes causal sense of a whole series of features of the planetary motions is more likely to be true than one that leaves these features unexplained.

Copernicus and those who followed him believed that they had good arguments for the reality of the earth's motion around the sun. They sometimes overstated the force of those arguments, to be sure, using terms like 'proof' and 'demonstration'. The natural philosophers of the day were not yet accustomed to the weaker notions of likelihood and probability. Galileo found, to his cost, that he had to speak in terms of demonstration if his claims for the Copernican system were to be taken seriously. He did not have a demonstration, but from our perspective, he called effectively on the criterion of coherence in his critique of the geostatic alternative, just as Copernicus had earlier done.

As we look back on those debates, we are ready to allow that the coherence arguments of Copernicus and Galileo *did* carry force, that they *did* give a motive for accepting the new heliocentric model as true. And their force came from something other than predictive advantage. Kuhn's point in regard to theory assessment, one that became clearer in his successive formulations of it, was that the different theory values were not reducible to one another, and hence

that no simple algorithm, no logic of confirmation such as the logical positivists had sought, underlay real-life theory decision. What I have tried to do here is to carry this insight further and to note the special epistemic weight carried by certain of these values. Besides coherence, one could make similar cases for fertility and unifying power. It is hard to make sense of the role played by these values if one adopts the instrumentalist standpoint that Kuhn feels compelled to advocate.

The case for scientific realism rests in large part on these "superempirical" values. That is, when we ask about a particular theory, how likely is it that it is true (correlatively, how likely is it that something like the explanatory entities it postulates actually exist), it is to these virtues that we are inclined to turn. To say that a theory simply "saves the phenomena," though this carries *some* epistemic weight, leaves open the suspicion of its being ad hoc. If a theory be thought of simply as an hypothetico-deductive device, it would seem plausible to suppose that other devices might account as well or better for the phenomena to be explained. It is only when the *temporal* dimension is added, when a theory is evaluated in a historical context, when its success in unifying domains over time or in predicting new sorts of phenomena are taken into account, that conviction begins to emerge. Theories are not assessed simply as predictors; they are not confirmed purely by the enumeration of consequences.

My conclusion is that the diversity of the expectations scientists hold up for their theories argues not only for the tentative character of theory choice, Kuhn's original point, but also for its properly epistemic character. This leaves us, of course, with a problem: how can the difficulties in regard to incommensurability be reconciled with the epistemic force of such arguments as that of Copernicus? Kuhn emphasized the discontinuities of language across theory change so strongly that he left no room for the possibility of convergence, for the possibility that the theories of the paleontologists of today, for example, not only solve more puzzles than those of yesteryear but also tell us, with high degree of likelihood, what actually happened at distant epochs in the earth's past.

The Kuhnian heritage is thus a curiously divided one. Kuhn wanted to maintain the rational character of theory choice in science while denying the epistemic character of the theory chosen. The consequent tensions are, of course, familiar to every reader of current philosophy of science. Thirty years later, *The Structure of Scientific Revolutions* still leaves us with an agenda.

Notes

1. See *The Essential Tension (ET),* pp. 320–339. In his effort to ward off the charge of subjectivism, Kuhn might also have pointed to "The Function of Measurement in Modern Physical Science" (*Isis* 52 [1961]: 161–190; reprinted in *ET,* pp. 178–224), which appeared before *SSR* and whose theme was that "measurement can be an immensely powerful weapon in the battle between two theories" (*ET,* p. 211), that "the comparison of numerical predictions . . . has proved particularly successful in bringing scientific controversies to a close" (*ET,* p. 213). Or he could have recalled an even earlier paper, "The Essential Tension" (*The Third University of Utah Research Conference on the Identification of Scientific Talent,* ed. C. W. Taylor [Salt Lake City: University of Utah Press, 1959], 162–174; reprinted in *ET,* pp. 225–239), whose title referred to the opposition between the themes of tradition and innovation in science and which argued that it is the very effort to work within a tightly construed tradition that leads eventually to the recognition of anomalies that in turn prepares the way for revolution (*ET,* p. 234). One further paper that Kuhn might have called on was "A Function for Thought Experiments" (*L'Aventure de la science,* ed. Mélanges Alexandre Koyré [Paris: Hermann, 1964], vol. 2, pp. 307–334; reprinted in *ET,* pp. 240–265), which describes how failures of expectation induce the crisis that is the usual prelude to paradigm change (*ET,* p. 263).

2. "Objectivity," p. 30.

3. *The Structure of Scientific Revolutions,* 2nd ed. (Chicago: University of Chicago Press, 1970), p. 199.

4. "Metaphor in Science," in *Metaphor and Thought,* ed. Andrew Ortony (Cambridge: Cambridge University Press, 1979), 409–419; see p. 416.

5. In a recent essay Kuhn distinguishes between translation and interpretation and shows how communication can occur even where languages are incommensurable ("Commensurability, Comparability, Communicability," *PSA 1982* [Philosophy of Science Association], 1983: 669–688). In a comment Philip Kitcher remarks that Kuhn, in his later readings of *SSR,* has progressively weakened the dramatic doctrine of the original work in ways, be it said, of which Kitcher approves ("Implications of Incommensurability," *PSA 1982,* 1983: 689–703).

6. *SSR,* p. 204.

7. *SSR,* p. 204.

8. "Objectivity," p. 324.

9. "Objectivity," p. 325.

10. *SSR,* p. 94.

11. *SSR,* p. 94.

12. In *SSR* Kuhn himself distinguishes between "major paradigm changes, like those attributable to Copernicus and Lavoisier," and "the far smaller ones associated with the assimilation of a new sort of phenomenon, like oxygen or X-rays" (p. 92).

13. *SSR,* p. 58.

14. *SSR,* p. 62. It is not clear to me that the discovery of X rays satisfies either of these criteria in any other than a minimal way.

15. *SSR,* p. 199.

16. *SSR,* p. 62. There might be some question as to whether, in fact, a single theory of electricity did emerge at this time. But that is not to the point of my inquiry.

17. I have worked out this theme in some detail in my "Conceptions of Science in the Scientific Revolution," in *Reappraisals of the Scientific Revolution,* ed. David Lindberg and Robert Westman (Cambridge: Cambridge University Press, 1990).

18. *SSR,* p. 182.

19. "Objectivity," p. 335.

20. See Larry Laudan, *Science and Values* (Berkeley: University of California Press, 1984); Dudley Shapere, *Reason and the Search for Knowledge* (Dordrecht: Reidel, 1984). I have discussed the ironies of this particular divergence more fully in "The Shaping of Scientific Rationality," in *Construction and Constraint,* ed. E. McMullin (Notre Dame: University of Notre Dame Press, 1988), pp. 1–47.

21. Nicholas Rescher defends a somewhat similar position in regard to how different from ours the "science" carried on by the inhabitants of a distant planet might be: "Science *as we have it*—the only 'science' that we ourselves know—is a specifically human artifact that must be expected to reflect in significant degree the particular characteristics of its makers. Consequently, the prospect that an alien 'science'-possessing civilizaiton has a *science* that we could acknowledge (if sufficiently informed) as representing the same general line of inquiry as that in which we ourselves are engaged seems extremely implausible" ("Extraterrestrial Science," *Philosophia Naturalis* 21 [1984]: 400–424; see p. 413.

22. *SSR,* pp. 109–110.

23. "Objectivity," p. 336.

24. "Rationality and Theory Choice," *Journal of Philosophy* 80 (1983): 563–570; see p. 567.

25. Ibid., p. 568.

26. C. G. Hempel, "Valuation and Objectivity in Science," in *Physics, Philosophy, and Psychoanalysis,* ed. R. S. Cohen and L. Laudan (Dordrecht: Reidel, 1983), 73–100.

27. "Objectivity," p. 335.

28. See my "Values in Science," *PSA 1982,* 1983: 3–25.

29. The story sketched so lightly here is told in much more detail in my "Goals of Natural Science," *Proc. American Philosophical Association* 58 (1984): 37–64.

30. For a fuller historical treatment, see my "Conceptions of Science in the Scientific Revolution."

31. "Rationality and Theory Choice," p. 569.

32. *SSR,* p. 206.

33. *SSR,* p. 206.

34. Laudan's much-quoted essay, "A Confutation of Convergent Realism," in *Scientific Realism,* ed. J. Leplin (Berkeley: University of California Press, 1984), 218–249, presents in detail the sort of arguments that Kuhn would need to support his own rejection of convergence.

35. See my "Case for Scientific Realism," in *Scientific Realism,* ed. J. Leplin, 8–40, and "Selective Anti-realism," *Philosophical Studies* 61 (1991): 97–108.

36. *The Copernican Revolution* (New York: Random House, 1957), p. 171.

37. Ibid., p. 172.

38. Ibid., p. 178.

39. Ibid., p. 172.

40. Kepler's clearest treatment of this issue will be found in the *Apologia Tychonis contra Ursum* (1600). See Nicholas Jardine's translation of this work in *The Birth of History and Philosophy of Science* (Cambridge: Cambridge University Press, 1984). Michael Gardner extracts a "Kepler principle" to the effect that it counts in favor of the realistic acceptance of a theory if it explains facts that competing theories merely postulate. See "Realism and Instrumentalism in Pre-Newtonian Astronomy," in *Testing Scientific Theories,* ed. John Earman (Minneapolis: University of Minnesota Press, 1983), 201–265; p. 256.

The Historians Look