Postscript—1969

It has now been almost seven years since this book was first published.[1] In the interim both the response of critics and my own further work have increased my understanding of a number of the issues it raises. On fundamentals my viewpoint is very nearly unchanged, but I now recognize aspects of its initial formulation that create gratuitous difficulties and misunderstandings. Since some of those misunderstandings have been my own, their elimination enables me to gain ground that should ultimately provide the basis for a new version of the book.[2] Meanwhile, I welcome the chance to sketch needed revisions, to comment on some reiterated criticisms, and to suggest directions in which my own thought is presently developing.[3]

Several of the key difficulties of my original text cluster about the concept of a paradigm, and my discussion begins with them.[4] In the subsection that follows at once, I suggest the desirability of disentangling that concept from the notion of a scientific community, indicate how this may be done, and discuss some signifi-

[1] This postscript was first prepared at the suggestion of my onetime student and longtime friend, Dr. Shigeru Nakayama of the University of Tokyo, for inclusion in his Japanese translation of this book. I am grateful to him for the idea, for his patience in awaiting its fruition, and for permission to include the result in the English language edition.

[2] For this edition I have attempted no systematic rewriting, restricting alterations to a few typographical errors plus two passages which contained isolable errors. One of these is the description of the role of Newton's *Principia* in the development of eighteenth-century mechanics on pp. 30–33, above. The other concerns the response to crises on p. 84.

[3] Other indications will be found in two recent essays of mine: "Reflection on My Critics," in Imre Lakatos and Alan Musgrave (eds.), *Criticism and the Growth of Knowledge* (Cambridge, 1970); and "Second Thoughts on Paradigms," in Frederick Suppe (ed.), *The Structure of Scientific Theories* (Urbana, Ill., 1970 or 1971), both currently in press. I shall cite the first of these essays below as "Reflections" and the volume in which it appears as *Growth of Knowledge;* the second essay will be referred to as "Second Thoughts."

[4] For particularly cogent criticism of my initial presentation of paradigms see: Margaret Masterman, "The Nature of a Paradigm," in *Growth of Knowledge;* and Dudley Shapere, "The Structure of Scientific Revolutions," *Philosophical Review*, LXXIII (1964), 383–94.

cant consequences of the resulting analytic separation. Next I consider what occurs when paradigms are sought by examining the behavior of the members of a *previously determined* scientific community. That procedure quickly discloses that in much of the book the term 'paradigm' is used in two different senses. On the one hand, it stands for the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community. On the other, it denotes one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal science. The first sense of the term, call it the sociological, is the subject of Subsection 2, below; Subsection 3 is devoted to paradigms as exemplary past achievements.

Philosophically, at least, this second sense of 'paradigm' is the deeper of the two, and the claims I have made in its name are the main sources for the controversies and misunderstandings that the book has evoked, particularly for the charge that I make of science a subjective and irrational enterprise. These issues are considered in Subsections 4 and 5. The first argues that terms like 'subjective' and 'intuitive' cannot appropriately be applied to the components of knowledge that I have described as tacitly embedded in shared examples. Though such knowledge is not, without essential change, subject to paraphrase in terms of rules and criteria, it is nevertheless systematic, time tested, and in some sense corrigible. Subsection 5 applies that argument to the problem of choice between two incompatible theories, urging in brief conclusion that men who hold incommensurable viewpoints be thought of as members of different language communities and that their communication problems be analyzed as problems of translation. Three residual issues are discussed in the concluding Subsections, 6 and 7. The first considers the charge that the view of science developed in this book is through-and-through relativistic. The second begins by inquiring whether my argument really suffers, as has been said, from a confusion between the descriptive and the normative modes; it concludes with brief remarks on a topic deserving a separate essay: the extent to which the book's main theses may legitimately be applied to fields other than science.

1. *Paradigms and Community Structure*

The term 'paradigm' enters the preceding pages early, and its manner of entry is intrinsically circular. A paradigm is what the members of a scientific community share, *and,* conversely, a scientific community consists of men who share a paradigm. Not all circularities are vicious (I shall defend an argument of similar structure late in this postscript), but this one is a source of real difficulties. Scientific communities can and should be isolated without prior recourse to paradigms; the latter can then be discovered by scrutinizing the behavior of a given community's members. If this book were being rewritten, it would therefore open with a discussion of the community structure of science, a topic that has recently become a significant subject of sociological research and that historians of science are also beginning to take seriously. Preliminary results, many of them still unpublished, suggest that the empirical techniques required for its exploration are non-trivial, but some are in hand and others are sure to be developed.[5] Most practicing scientists respond at once to questions about their community affiliations, taking for granted that responsibility for the various current specialties is distributed among groups of at least roughly determinate membership. I shall therefore here assume that more systematic means for their identification will be found. Instead of presenting preliminary research results, let me briefly articulate the intuitive notion of community that underlies much in the earlier chapters of this book. It is a notion now widely shared by scientists, sociologists, and a number of historians of science.

[5] W. O. Hagstrom, *The Scientific Community* (New York, 1965), chaps. iv and v; D. J. Price and D. de B. Beaver, "Collaboration in an Invisible College," *American Psychologist*, XXI (1966), 1011–18; Diana Crane, "Social Structure in a Group of Scientists: A Test of the 'Invisible College' Hypothesis," *American Sociological Review*, XXXIV (1969), 335–52; N. C. Mullins, *Social Networks among Biological Scientists*, (Ph.D. diss., Harvard University, 1966), and "The Micro-Structure of an Invisible College: The Phage Group" (paper delivered at an annual meeting of the American Sociological Association, Boston, 1968).

A scientific community consists, on this view, of the practitioners of a scientific specialty. To an extent unparalleled in most other fields, they have undergone similar educations and professional initiations; in the process they have absorbed the same technical literature and drawn many of the same lessons from it. Usually the boundaries of that standard literature mark the limits of a scientific subject matter, and each community ordinarily has a subject matter of its own. There are schools in the sciences, communities, that is, which approach the same subject from incompatible viewpoints. But they are far rarer there than in other fields; they are always in competition; and their competition is usually quickly ended. As a result, the members of a scientific community see themselves and are seen by others as the men uniquely responsible for the pursuit of a set of shared goals, including the training of their successors. Within such groups communication is relatively full and professional judgment relatively unanimous. Because the attention of different scientific communities is, on the other hand, focused on different matters, professional communication across group lines is sometimes arduous, often results in misunderstanding, and may, if pursued, evoke significant and previously unsuspected disagreement.

Communities in this sense exist, of course, at numerous levels. The most global is the community of all natural scientists. At an only slightly lower level the main scientific professional groups are communities: physicists, chemists, astronomers, zoologists, and the like. For these major groupings, community membership is readily established except at the fringes. Subject of highest degree, membership in professional societies, and journals read are ordinarily more than sufficient. Similar techniques will also isolate major subgroups: organic chemists, and perhaps protein chemists among them, solid-state and high-energy physicists, radio astronomers, and so on. It is only at the next lower level that empirical problems emerge. How, to take a contemporary example, would one have isolated the phage group prior to its public acclaim? For this purpose one must have recourse to attendance at special conferences, to the distribution of draft manuscripts or galley proofs prior to publication, and above all to formal and informal communication networks including those discovered in correspondence and in the linkages among citations.[6] I take it that the job can and will be done, at least for the contemporary scene and the more recent parts of the historical. Typically it may yield communities of perhaps one hundred members, occasionally significantly fewer. Usually individual scientists, particularly the ablest, will belong to several such groups either simultaneously or in succession.

Communities of this sort are the units that this book has presented as the producers and validators of scientific knowledge. Paradigms are something shared by the members of such groups. Without reference to the nature of these shared elements, many aspects of science described in the preceding pages can scarcely be understood. But other aspects can, though they are not independently presented in my original text. It is therefore worth noting, before turning to paradigms directly, a series of issues that require reference to community structure alone.

Probably the most striking of these is what I have previously called the transition from the pre- to the post-paradigm period in the development of a scientific field. That transition is the one sketched above in Section II. Before it occurs, a number of schools compete for the domination of a given field. Afterward, in the wake of some notable scientific achievement, the number of schools is greatly reduced, ordinarily to one, and a more efficient mode of scientific practice begins. The latter is generally esoteric and oriented to puzzle-solving, as the work of a group can be only when its members take the foundations of their field for granted.

The nature of that transition to maturity deserves fuller discussion than it has received in this book, particularly from those concerned with the development of the contemporary social

[6] Eugene Garfield, *The Use of Citation Data in Writing the History of Science* (Philadelphia: Institute of Scientific Information, 1964); M. M. Kessler, "Comparison of the Results of Bibliographic Coupling and Analytic Subject Indexing," *American Documentation*, XVI (1965), 223–33; D. J. Price, "Networks of Scientific Papers," *Science*, CIL (1965), 510–15.

sciences. To that end it may help to point out that the transition need not (I now think should not) be associated with the first acquisition of a paradigm. The members of all scientific communities, including the schools of the "pre-paradigm" period, share the sorts of elements which I have collectively labelled 'a paradigm.' What changes with the transition to maturity is not the presence of a paradigm but rather its nature. Only after the change is normal puzzle-solving research possible. Many of the attributes of a developed science which I have above associated with the acquisition of a paradigm I would therefore now discuss as consequences of the acquisition of the sort of paradigm that identifies challenging puzzles, supplies clues to their solution, and guarantees that the truly clever practitioner will succeed. Only those who have taken courage from observing that their own field (or school) has paradigms are likely to feel that something important is sacrificed by the change.

A second issue, more important at least to historians, concerns this book's implicit one-to-one identification of scientific communities with scientific subject matters. I have, that is, repeatedly acted as though, say, 'physical optics,' 'electricity,' and 'heat' must name scientific communities because they do name subject matters for research. The only alternative my text has seemed to allow is that all these subjects have belonged to the physics community. Identifications of that sort will not, however, usually withstand examination, as my colleagues in history have repeatedly pointed out. There was, for example, no physics community before the mid-nineteenth century, and it was then formed by the merger of parts of two previously separate communities, mathematics and natural philosophy *(physique expérimentale)*. What is today the subject matter for a single broad community has been variously distributed among diverse communities in the past. Other narrower subjects, for example heat and the theory of matter, have existed for long periods without becoming the special province of any single scientific community. Both normal science and revolutions are, however, community-based activities. To discover and analyze them, one must first unravel the changing community structure of the sciences over time. A paradigm governs, in the first instance, not a subject matter but rather a group of practitioners. Any study of paradigm-directed or of paradigm-shattering research must begin by locating the responsible group or groups.

When the analysis of scientific development is approached in that way, several difficulties which have been foci for critical attention are likely to vanish. A number of commentators have, for example, used the theory of matter to suggest that I drastically overstate the unanimity of scientists in their allegiance to a paradigm. Until comparatively recently, they point out, those theories have been topics for continuing disagreement and debate. I agree with the description but think it no counter-example. Theories of matter were not, at least until about 1920, the special province or the subject matter for any scientific community. Instead, they were tools for a large number of specialists' groups. Members of different communities sometimes chose different tools and criticized the choice made by others. Even more important, a theory of matter is not the sort of topic on which the members of even a single community must necessarily agree. The need for agreement depends on what it is the community does. Chemistry in the first half of the nineteenth century provides a case in point. Though several of the community's fundamental tools—constant proportion, multiple proportion, and combining weights—had become common property as a result of Dalton's atomic theory, it was quite possible for chemists, after the event, to base their work on these tools and to disagree, sometimes vehemently, about the existence of atoms.

Some other difficulties and misunderstandings will, I believe, be dissolved in the same way. Partly because of the examples I have chosen and partly because of my vagueness about the nature and size of the relevant communities, a few readers of this book have concluded that my concern is primarily or exclusively with major revolutions such as those associated with Copernicus, Newton, Darwin, or Einstein. A clearer delineation of community structure should, however, help to enforce the rather different impression I have tried to create. A revolution

is for me a special sort of change involving a certain sort of reconstruction of group commitments. But it need not be a large change, nor need it seem revolutionary to those outside a single community, consisting perhaps of fewer than twenty-five people. It is just because this type of change, little recognized or discussed in the literature of the philosophy of science, occurs so regularly on this smaller scale that revolutionary, as against cumulative, change so badly needs to be understood.

One last alteration, closely related to the preceding, may help to facilitate that understanding. A number of critics have doubted whether crisis, the common awareness that something has gone wrong, precedes revolutions so invariably as I have implied in my original text. Nothing important to my argument depends, however, on crises' being an absolute prerequisite to revolutions; they need only be the usual prelude, supplying, that is, a self-correcting mechanism which ensures that the rigidity of normal science will not forever go unchallenged. Revolutions may also be induced in other ways, though I think they seldom are. In addition, I would now point out what the absence of an adequate discussion of community structure has obscured above: crises need not be generated by the work of the community that experiences them and that sometimes undergoes revolution as a result. New instruments like the electron microscope or new laws like Maxwell's may develop in one specialty and their assimilation create crisis in another.

2. *Paradigms as the Constellation of Group Commitments*

Turn now to paradigms and ask what they can possibly be. My original text leaves no more obscure or important question. One sympathetic reader, who shares my conviction that 'paradigm' names the central philosophical elements of the book, prepared a partial analytic index and concluded that the term is used in at least twenty-two different ways.[7] Most of those differences are, I now think, due to stylistic inconsistencies (e.g., Newton's Laws are sometimes a paradigm, sometimes parts of a paradigm, and

[7] Masterman, *op. cit.*

sometimes paradigmatic), and they can be eliminated with relative ease. But, with that editorial work done, two very different usages of the term would remain, and they require separation. The more global use is the subject of this subsection; the other will be considered in the next.

Having isolated a particular community of specialists by techniques like those just discussed, one may usefully ask: What do its members share that accounts for the relative fulness of their professional communication and the relative unanimity of their professional judgments? To that question my original text licenses the answer, a paradigm or set of paradigms. But for this use, unlike the one to be discussed below, the term is inappropriate. Scientists themselves would say they share a theory or set of theories, and I shall be glad if the term can ultimately be recaptured for this use. As currently used in philosophy of science, however, 'theory' connotes a structure far more limited in nature and scope than the one required here. Until the term can be freed from its current implications, it will avoid confusion to adopt another. For present purposes I suggest 'disciplinary matrix': 'disciplinary' because it refers to the common possession of the practitioners of a particular discipline; 'matrix' because it is composed of ordered elements of various sorts, each requiring further specification. All or most of the objects of group commitment that my original text makes paradigms, parts of paradigms, or paradigmatic are constituents of the disciplinary matrix, and as such they form a whole and function together. They are, however, no longer to be discussed as though they were all of a piece. I shall not here attempt an exhaustive list, but noting the main sorts of components of a disciplinary matrix will both clarify the nature of my present approach and simultaneously prepare for my next main point.

One important sort of component I shall label 'symbolic generalizations,' having in mind those expressions, deployed without question or dissent by group members, which can readily be cast in a logical form like $(x)(y)(z)\phi(x, y, z)$. They are the formal or the readily formalizable components of the disciplinary matrix. Sometimes they are found already in symbolic form: $f = ma$ or $I = V/R$. Others are ordinarily expressed in words: "elements combine in constant proportion by weight," or "action equals reaction." If it were not for the general acceptance of expressions like these, there would be no points at which group members could attach the powerful techniques of logical and mathematical manipulation in their puzzle-solving enterprise. Though the example of taxonomy suggests that normal science can proceed with few such expressions, the power of a science seems quite generally to increase with the number of symbolic generalizations its practioners have at their disposal.

These generalizations look like laws of nature, but their function for group members is not often that alone. Sometimes it is: for example the Joule-Lenz Law, $H = RI^2$. When that law was discovered, community members already knew what H, R, and I stood for, and these generalizations simply told them something about the behavior of heat, current, and resistance that they had not known before. But more often, as discussion earlier in the book indicates, symbolic generalizations simultaneously serve a second function, one that is ordinarily sharply separated in analyses by philosophers of science. Like $f = ma$ or $I = V/R$, they function in part as laws but also in part as definitions of some of the symbols they deploy. Furthermore, the balance between their inseparable legislative and definitional force shifts over time. In another context these points would repay detailed analysis, for the nature of the commitment to a law is very different from that of commitment to a definition. Laws are often corrigible piecemeal, but definitions, being tautologies, are not. For example, part of what the acceptance of Ohm's Law demanded was a redefinition of both 'current' and 'resistance'; if those terms had continued to mean what they had meant before, Ohm's Law could not have been right; that is why it was so strenuously opposed as, say, the Joule-Lenz Law was not.[8] Probably that situation is typical. I currently suspect that

[8] For significant parts of this episode see: T. M. Brown, "The Electric Current in Early Nineteenth-Century French Physics," *Historical Studies in the Physical Sciences*, I (1969), 61–103, and Morton Schagrin, "Resistance to Ohm's Law," *American Journal of Physics*, XXI (1963), 536–47.

all revolutions involve, among other things, the abandonment of generalizations the force of which had previously been in some part that of tautologies. Did Einstein show that simultaneity was relative or did he alter the notion of simultaneity itself? Were those who heard paradox in the phrase 'relativity of simultaneity' simply wrong?

Consider next a second type of component of the disciplinary matrix, one about which a good deal has been said in my original text under such rubrics as 'metaphysical paradigms' or 'the metaphysical parts of paradigms.' I have in mind shared commitments to such beliefs as: heat is the kinetic energy of the constituent parts of bodies; all perceptible phenomena are due to the interaction of qualitatively neutral atoms in the void, or, alternatively, to matter and force, or to fields. Rewriting the book now I would describe such commitments as beliefs in particular models, and I would expand the category models to include also the relatively heuristic variety: the electric circuit may be regarded as a steady-state hydrodynamic system; the molecules of a gas behave like tiny elastic billiard balls in random motion. Though the strength of group commitment varies, with non-trivial consequences, along the spectrum from heuristic to ontological models, all models have similar functions. Among other things they supply the group with preferred or permissible analogies and metaphors. By doing so they help to determine what will be accepted as an explanation and as a puzzle-solution; conversely, they assist in the determination of the roster of unsolved puzzles and in the evaluation of the importance of each. Note, however, that the members of scientific communities may not have to share even heuristic models, though they usually do so. I have already pointed out that membership in the community of chemists during the first half of the nineteenth century did not demand a belief in atoms.

A third sort of element in the disciplinary matrix I shall here describe as values. Usually they are more widely shared among different communities than either symbolic generalizations or models, and they do much to provide a sense of community to natural scientists as a whole. Though they function at all times, their particular importance emerges when the members of a

particular community must identify crisis or, later, choose between incompatible ways of practicing their discipline. Probably the most deeply held values concern predictions: they should be accurate; quantitative predictions are preferable to qualitative ones; whatever the margin of permissible error, it should be consistently satisfied in a given field; and so on. There are also, however, values to be used in judging whole theories: they must, first and foremost, permit puzzle-formulation and solution; where possible they should be simple, self-consistent, and plausible, compatible, that is, with other theories currently deployed. (I now think it a weakness of my original text that so little attention is given to such values as internal and external consistency in considering sources of crisis and factors in theory choice.) Other sorts of values exist as well—for example, science should (or need not) be socially useful—but the preceding should indicate what I have in mind.

One aspect of shared values does, however, require particular mention. To a greater extent than other sorts of components of the disciplinary matrix, values may be shared by men who differ in their application. Judgments of accuracy are relatively, though not entirely, stable from one time to another and from one member to another in a particular group. But judgments of simplicity, consistency, plausibility, and so on often vary greatly from individual to individual. What was for Einstein an insupportable inconsistency in the old quantum theory, one that rendered the pursuit of normal science impossible, was for Bohr and others a difficulty that could be expected to work itself out by normal means. Even more important, in those situations where values must be applied, different values, taken alone, would often dictate different choices. One theory may be more accurate but less consistent or plausible than another; again the old quantum theory provides an example. In short, though values are widely shared by scientists and though commitment to them is both deep and constitutive of science, the application of values is sometimes considerably affected by the features of individual personality and biography that differentiate the members of the group.

To many readers of the preceding chapters, this characteristic of the operation of shared values has seemed a major weakness of my position. Because I insist that what scientists share is not sufficient to command uniform assent about such matters as the choice between competing theories or the distinction between an ordinary anomaly and a crisis-provoking one, I am occasionally accused of glorifying subjectivity and even irrationality.[9] But that reaction ignores two characteristics displayed by value judgments in any field. First, shared values can be important determinants of group behavior even though the members of the group do not all apply them in the same way. (If that were not the case, there would be no *special* philosophic problems about value theory or aesthetics.) Men did not all paint alike during the periods when representation was a primary value, but the developmental pattern of the plastic arts changed drastically when that value was abandoned.[10] Imagine what would happen in the sciences if consistency ceased to be a primary value. Second, individual variability in the application of shared values may serve functions essential to science. The points at which values must be applied are invariably also those at which risks must be taken. Most anomalies are resolved by normal means; most proposals for new theories do prove to be wrong. If all members of a community responded to each anomaly as a source of crisis or embraced each new theory advanced by a colleague, science would cease. If, on the other hand, no one reacted to anomalies or tc brand-new theories in high-risk ways, there would be few or no revolutions. In matters like these the resort to shared values rather than to shared rules governing individual choice may be the community's way of distributing risk and assuring the long-term success of its enterprise.

Turn now to a fourth sort of element in the disciplinary matrix, not the only other kind but the last I shall discuss here. For it the term 'paradigm' would be entirely appropriate, both philologi-

[9] See particularly: Dudley Shapere, "Meaning and Scientific Change," in *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, The University of Pittsburgh Series in the Philosophy of Science, III (Pittsburgh, 1966), 41–85; Israel Scheffler, *Science and Subjectivity* (New York, 1967); and the essays of Sir Karl Popper and Imre Lakatos in *Growth of Knowledge*.

[10] See the discussion at the beginning of Section XIII, above.

cally and autobiographically; this is the component of a group's shared commitments which first led me to the choice of that word. Because the term has assumed a life of its own, however, I shall here substitute 'exemplars.' By it I mean, initially, the concrete problem-solutions that students encounter from the start of their scientific education, whether in laboratories, on examinations, or at the ends of chapters in science texts. To these shared examples should, however, be added at least some of the technical problem-solutions found in the periodical literature that scientists encounter during their post-educational research careers and that also show them by example how their job is to be done. More than other sorts of components of the disciplinary matrix, differences between sets of exemplars provide the community fine-structure of science. All physicists, for example, begin by learning the same exemplars: problems such as the inclined plane, the conical pendulum, and Keplerian orbits; instruments such as the vernier, the calorimeter, and the Wheatstone bridge. As their training develops, however, the symbolic generalizations they share are increasingly illustrated by different exemplars. Though both solid-state and field-theoretic physicists share the Schrödinger equation, only its more elementary applications are common to both groups.

3. *Paradigms as Shared Examples*

The paradigm as shared example is the central element of what I now take to be the most novel and least understood aspect of this book. Exemplars will therefore require more attention than the other sorts of components of the disciplinary matrix. Philosophers of science have not ordinarily discussed the problems encountered by a student in laboratories or in science texts, for these are thought to supply only practice in the application of what the student already knows. He cannot, it is said, solve problems at all unless he has first learned the theory and some rules for applying it. Scientific knowledge is embedded in theory and rules; problems are supplied to gain facility in their application. I have tried to argue, however, that this localization of the cognitive content of science is wrong. After the student has done many problems, he may gain only added facility by solving more. But at the start and for some time after, doing problems is learning consequential things about nature. In the absence of such exemplars, the laws and theories he has previously learned would have little empirical content.

To indicate what I have in mind I revert briefly to symbolic generalizations. One widely shared example is Newton's Second Law of Motion, generally written as $f = ma$. The sociologist, say, or the linguist who discovers that the corresponding expression is unproblematically uttered and received by the members of a given community will not, without much additional investigation, have learned a great deal about what either the expression or the terms in it mean, about how the scientists of the community attach the expression to nature. Indeed, the fact that they accept it without question and use it as a point at which to introduce logical and mathematical manipulation does not of itself imply that they agree at all about such matters as meaning and application. Of course they do agree to a considerable extent, or the fact would rapidly emerge from their subsequent conversation. But one may well ask at what point and by what means they have come to do so. How have they learned, faced with a given experimental situation, to pick out the relevant forces, masses, and accelerations?

In practice, though this aspect of the situation is seldom or never noted, what students have to learn is even more complex than that. It is not quite the case that logical and mathematical manipulation are applied directly to $f = ma$. That expression proves on examination to be a law-sketch or a law-schema. As the student or the practicing scientist moves from one problem situation to the next, the symbolic generalization to which such manipulations apply changes. For the case of free fall, $f = ma$ becomes $mg = m\frac{d^2s}{dt^2}$; for the simple pendulum it is transformed to $mg \sin\theta = -ml\frac{d^2\theta}{dt^2}$; for a pair of interacting harmonic oscillators it becomes two equations, the first of which may be written

$m_1\frac{d^2s_1}{dt^2} + k_1s_1 = k_2(s_2 - s_1 + d)$; and for more complex situations, such as the gyroscope, it takes still other forms, the family resemblance of which to $f = ma$ is still harder to discover. Yet, while learning to identify forces, masses, and accelerations in a variety of physical situations not previously encountered, the student has also learned to design the appropriate version of $f = ma$ through which to interrelate them, often a version for which he has encountered no literal equivalent before. How has he learned to do this?

A phenomenon familiar to both students of science and historians of science provides a clue. The former regularly report that they have read through a chapter of their text, understood it perfectly, but nonetheless had difficulty solving a number of the problems at the chapter's end. Ordinarily, also, those difficulties dissolve in the same way. The student discovers, with or without the assistance of his instructor, a way to see his problem as *like* a problem he has already encountered. Having seen the resemblance, grasped the analogy between two or more distinct problems, he can interrelate symbols and attach them to nature in the ways that have proved effective before. The law-sketch, say $f = ma$, has functioned as a tool, informing the student what similarities to look for, signaling the gestalt in which the situation is to be seen. The resultant ability to see a variety of situations as like each other, as subjects for $f = ma$ or some other symbolic generalization, is, I think, the main thing a student acquires by doing exemplary problems, whether with a pencil and paper or in a well-designed laboratory. After he has completed a certain number, which may vary widely from one individual to the next, he views the situations that confront him as a scientist in the same gestalt as other members of his specialists' group. For him they are no longer the same situations he had encountered when his training began. He has meanwhile assimilated a time-tested and group-licensed way of seeing.

The role of acquired similarity relations also shows clearly in the history of science. Scientists solve puzzles by modeling them on previous puzzle-solutions, often with only minimal recourse to symbolic generalizations. Galileo found that a ball rolling down an incline acquires just enough velocity to return it to the same vertical height on a second incline of any slope, and he learned to see that experimental situation as like the pendulum with a point-mass for a bob. Huyghens then solved the problem of the center of oscillation of a physical pendulum by imagining that the extended body of the latter was composed of Galilean point-pendula, the bonds between which could be instantaneously released at any point in the swing. After the bonds were released, the individual point-pendula would swing freely, but their collective center of gravity when each attained its highest point would, like that of Galileo's pendulum, rise only to the height from which the center of gravity of the extended pendulum had begun to fall. Finally, Daniel Bernoulli discovered how to make the flow of water from an orifice resemble Huyghens' pendulum. Determine the descent of the center of gravity of the water in tank and jet during an infinitesimal interval of time. Next imagine that each particle of water afterward moves separately upward to the maximum height attainable with the velocity acquired during that interval. The ascent of the center of gravity of the individual particles must then equal the descent of the center of gravity of the water in tank and jet. From that view of the problem the long-sought speed of efflux followed at once.[11]

That example should begin to make clear what I mean by learning from problems to see situations as like each other, as subjects for the application of the same scientific law or law-sketch. Simultaneously it should show why I refer to the consequential knowledge of nature acquired while learning the similarity relationship and thereafter embodied in a way of viewing

[11] For the example, see: René Dugas, *A History of Mechanics*, trans. J. R. Maddox (Neuchatel, 1955), pp. 135–36, 186–93, and Daniel Bernoulli, *Hydrodynamica, sive de viribus et motibus fluidorum, commentarii opus academicum* (Strasbourg, 1738), Sec. iii. For the extent to which mechanics progressed during the first half of the eighteenth century by modelling one problem-solution on another, see Clifford Truesdell, "Reactions of Late Baroque Mechanics to Success, Conjecture, Error, and Failure in Newton's *Principia*," *Texas Quarterly*, X (1967), 238–58.

physical situations rather than in rules or laws. The three problems in the example, all of them exemplars for eighteenth-century mechanicians, deploy only one law of nature. Known as the Principle of *vis viva*, it was usually stated as: "Actual descent equals potential ascent." Bernoulli's application of the law should suggest how consequential it was. Yet the verbal statement of the law, taken by itself, is virtually impotent. Present it to a contemporary student of physics, who knows the words and can do all these problems but now employs different means. Then imagine what the words, though all well known, can have said to a man who did not know even the problems. For him the generalization could begin to function only when he learned to recognize "actual descents" and "potential ascents" as ingredients of nature, and that is to learn something, prior to the law, about the situations that nature does and does not present. That sort of learning is not acquired by exclusively verbal means. Rather it comes as one is given words together with concrete examples of how they function in use; nature and words are learned together. To borrow once more Michael Polanyi's useful phrase, what results from this process is "tacit knowledge" which is learned by doing science rather than by acquiring rules for doing it.

4. *Tacit Knowledge and Intuition*

That reference to tacit knowledge and the concurrent rejection of rules isolates another problem that has bothered many of my critics and seemed to provide a basis for charges of subjectivity and irrationality. Some readers have felt that I was trying to make science rest on unanalyzable individual intuitions rather than on logic and law. But that interpretation goes astray in two essential respects. First, if I am talking at all about intuitions, they are not individual. Rather they are the tested and shared possessions of the members of a successful group, and the novice acquires them through training as a part of his preparation for group-membership. Second, they are not in principle unanalyzable. On the contrary, I am currently experimenting with a computer program designed to investigate their properties at an elementary level.

About that program I shall have nothing to say here,[12] but even mention of it should make my most essential point. When I speak of knowledge embedded in shared exemplars, I am not referring to a mode of knowing that is less systematic or less analyzable than knowledge embedded in rules, laws, or criteria of identification. Instead I have in mind a manner of knowing which is miscontrued if reconstructed in terms of rules that are first abstracted from exemplars and thereafter function in their stead. Or, to put the same point differently, when I speak of acquiring from exemplars the ability to recognize a given situation as like some and unlike others that one has seen before, I am not suggesting a process that is not potentially fully explicable in terms of neuro-cerebral mechanism. Instead I am claiming that the explication will not, by its nature, answer the question, "Similar with respect to what?" That question is a request for a rule, in this case for the criteria by which particular situations are grouped into similarity sets, and I am arguing that the temptation to seek criteria (or at least a full set) should be resisted in this case. It is not, however, system but a particular sort of system that I am opposing.

To give that point substance, I must briefly digress. What follows seems obvious to me now, but the constant recourse in my original text to phrases like "the world changes" suggests that it has not always been so. If two people stand at the same place and gaze in the same direction, we must, under pain of solipsism, conclude that they receive closely similar stimuli. (If both could put their eyes at the same place, the stimuli would be identical.) But people do not see stimuli; our knowledge of them is highly theoretical and abstract. Instead they have sensations, and we are under no compulsion to suppose that the sensations of our two viewers are the same. (Sceptics might remember that color blindness was nowhere noticed until John Dalton's description of it in 1794.) On the contrary, much

[12] Some information on this subject can be found in "Second Thoughts."

neural processing takes place between the receipt of a stimulus and the awareness of a sensation. Among the few things that we know about it with assurance are: that very different stimuli can produce the same sensations; that the same stimulus can produce very different sensations; and, finally, that the route from stimulus to sensation is in part conditioned by education. Individuals raised in different societies behave on some occasions as though they saw different things. If we were not tempted to identify stimuli one-to-one with sensations, we might recognize that they actually do so.

Notice now that two groups, the members of which have systematically different sensations on receipt of the same stimuli, do *in some sense* live in different worlds. We posit the existence of stimuli to explain our perceptions of the world, and we posit their immutability to avoid both individual and social solipsism. About neither posit have I the slightest reservation. But our world is populated in the first instance not by stimuli but by the objects of our sensations, and these need not be the same, individual to individual or group to group. To the extent, of course, that individuals belong to the same group and thus share education, language, experience, and culture, we have good reason to suppose that their sensations are the same. How else are we to understand the fulness of their communication and the communality of their behavioral responses to their environment? They must see things, process stimuli, in much the same ways. But where the differentiation and specialization of groups begins, we have no similar evidence for the immutability of sensation. Mere parochialism, I suspect, makes us suppose that the route from stimuli to sensation is the same for the members of all groups.

Returning now to exemplars and rules, what I have been trying to suggest, in however preliminary a fashion, is this. One of the fundamental techniques by which the members of a group, whether an entire culture or a specialists' sub-community within it, learn to see the same things when confronted with the same stimuli is by being shown examples of situations that their predecessors in the group have already learned to see as like each other and as different from other sorts of situations. These similar situations may be successive sensory presentations of the same individual—say of mother, who is ultimately recognized on sight as what she is and as different from father or sister. They may be presentations of the members of natural families, say of swans on the one hand and of geese on the other. Or they may, for the members of more specialized groups, be examples of the Newtonian situation, of situations, that is, that are alike in being subject to a version of the symbolic form $f = ma$ and that are different from those situations to which, for example, the law-sketches of optics apply.

Grant for the moment that something of this sort does occur. Ought we say that what has been acquired from exemplars is rules and the ability to apply them? That description is tempting because our seeing a situation as like ones we have encountered before must be the result of neural processing, fully governed by physical and chemical laws. In this sense, once we have learned to do it, recognition of similarity must be as fully systematic as the beating of our hearts. But that very parallel suggests that recognition may also be involuntary, a process over which we have no control. If it is, then we may not properly conceive it as something we manage by applying rules and criteria. To speak of it in those terms implies that we have access to alternatives, that we might, for example, have disobeyed a rule, or misapplied a criterion, or experimented with some other way of seeing.[13] Those, I take it, are just the sorts of things we cannot do.

Or, more precisely, those are things we cannot do until after we have had a sensation, perceived something. Then we do often seek criteria and put them to use. Then we may engage in interpretation, a deliberative process by which we choose among alternatives as we do not in perception itself. Perhaps, for example, something is odd about what we have seen (remember the anomalous playing cards). Turning a corner we see mother

[13] This point might never have needed making if all laws were like Newton's and all rules like the Ten Commandments. In that case the phrase 'breaking a law' would be nonsense, and a rejection of rules would not seem to imply a process not governed by law. Unfortunately, traffic laws and similar products of legislation can be broken, which makes the confusion easy.

entering a downtown store at a time we had thought she was home. Contemplating what we have seen we suddenly exclaim, "That wasn't mother, for she has red hair!" Entering the store we see the woman again and cannot understand how she could have been taken for mother. Or, perhaps we see the tail feathers of a waterfowl feeding from the bottom of a shallow pool. Is it a swan or a goose? We contemplate what we have seen, mentally comparing the tail feathers with those of swans and geese we have seen before. Or, perhaps, being proto-scientists, we simply want to know some general characteristic (the whiteness of swans, for example) of the members of a natural family we can already recognize with ease. Again, we contemplate what we have previously perceived, searching for what the members of the given family have in common.

These are all deliberative processes, and in them we do seek and deploy criteria and rules. We try, that is, to interpret sensations already at hand, to analyze what is for us the given. However we do that, the processes involved must ultimately be neural, and they are therefore governed by the same *physico-chemical* laws that govern perception on the one hand and the beating of our hearts on the other. But the fact that the system obeys the same laws in all three cases provides no reason to suppose that our neural apparatus is programmed to operate the same way in interpretation as in perception or in either as in the beating of our hearts. What I have been opposing in this book is therefore the attempt, traditional since Descartes but not before, to analyze perception as an interpretive process, as an unconscious version of what we do after we have perceived.

What makes the integrity of perception worth emphasizing is, of course, that so much past experience is embodied in the neural apparatus that transforms stimuli to sensations. An appropriately programmed perceptual mechanism has survival value. To say that the members of different groups may have different perceptions when confronted with the same stimuli is not to imply that they may have just any perceptions at all. In many environments a group that could not tell wolves from dogs could not endure. Nor would a group of nuclear physicists today survive as scientists if unable to recognize the tracks of alpha particles and electrons. It is just because so very few ways of seeing will do that the ones that have withstood the tests of group use are worth transmitting from generation to generation. Equally, it is because they have been selected for their success over historic time that we must speak of the experience and knowledge of nature embedded in the stimulus-to-sensation route.

Perhaps 'knowledge' is the wrong word, but there are reasons for employing it. What is built into the neural process that transforms stimuli to sensations has the following characteristics: it has been transmitted through education; it has, by trial, been found more effective than its historical competitors in a group's current environment; and, finally, it is subject to change both through further education and through the discovery of misfits with the environment. Those are characteristics of knowledge, and they explain why I use the term. But it is strange usage, for one other characteristic is missing. We have no direct access to what it is we know, no rules or generalizations with which to express this knowledge. Rules which could supply that access would refer to stimuli not sensations, and stimuli we can know only through elaborate theory. In its absence, the knowledge embedded in the stimulus-to-sensation route remains tacit.

Though it is obviously preliminary and need not be correct in all details, what has just been said about sensation is meant literally. At the very least it is a hypothesis about vision which should be subject to experimental investigation though probably not to direct check. But talk like this of seeing and sensation here also serves metaphorical functions as it does in the body of the book. We do not *see* electrons, but rather their tracks or else bubbles of vapor in a cloud chamber. We do not *see* electric currents at all, but rather the needle of an ammeter or galvanometer. Yet in the preceding pages, particularly in Section X, I have repeatedly acted as though we did perceive theoretical entities like currents, electrons, and fields, as though we learned to do so from examination of exemplars, and as though in these cases too it would be wrong to replace talk of seeing with talk of criteria and interpretation. The metaphor that transfers 'seeing' to contexts like these is scarcely a sufficient basis for such claims. In the long run it will need to be eliminated in favor of a more literal mode of discourse.

The computer program referred to above begins to suggest ways in which that may be done, but neither available space nor the extent of my present understanding permits my eliminating the metaphor here.[14] Instead I shall try briefly to bulwark it. Seeing water droplets or a needle against a numerical scale is a primitive perceptual experience for the man unacquainted with cloud chambers and ammeters. It thus requires contemplation, analysis, and interpretation (or else the intervention of external authority) before conclusions can be reached about electrons or currents. But the position of the man who has learned about these instruments and had much exemplary experience with them is very different, and there are corresponding differences in the way he processes the stimuli that reach him from them. Regarding the vapor in his breath on a cold winter afternoon, his sensation may be the same as that of a layman, but viewing a cloud chamber he sees (here literally) not droplets but the tracks of electrons, alpha particles, and so on. Those tracks are, if you will, criteria that he interprets as indices of the presence of the corresponding particles, but that route is both shorter and different from the one taken by the man who interprets droplets.

Or consider the scientist inspecting an ammeter to determine the number against which the needle has settled. His sensation probably is the same as the layman's, particularly if the latter has

[14] For readers of "Second Thoughts" the following cryptic remarks may be leading. The possibility of immediate recognition of the members of natural families depends upon the existence, after neural processing, of empty perceptual space between the families to be discriminated. If, for example, there were a perceived continuum of waterfowl ranging from geese to swans, we should be compelled to introduce a specific criterion for distinguishing them. A similar point can be made for unobservable entities. If a physical theory admits the existence of nothing else like an electric current, then a small number of criteria, which may vary considerably from case to case, will suffice to identify currents even though there is no set of rules that specifies the necessary and sufficient conditions for the identification. That point suggests a plausible corollary which may be more important. Given a set of necessary and sufficient conditions for identifying a theoretical entity, that entity can be eliminated from the ontology of a theory by substitution. In the absence of such rules, however, these entities are not eliminable; the theory then demands their existence.

read other sorts of meters before. But he has seen the meter (again often literally) in the context of the entire circuit, and he knows something about its internal structure. For him the needle's position is a criterion, but only of *the value* of the current. To interpret it he need determine only on which scale the meter is to be read. For the layman, on the other hand, the needle's position is not a criterion of anything except itself. To interpret it, he must examine the whole layout of wires, internal and external, experiment with batteries and magnets, and so on. In the metaphorical no less than in the literal use of 'seeing,' interpretation begins where perception ends. The two processes are not the same, and what perception leaves for interpretation to complete depends drastically on the nature and amount of prior experience and training.

5. *Exemplars, Incommensurability, and Revolutions*

What has just been said provides a basis for clarifying one more aspect of the book: my remarks on incommensurability and its consequences for scientists debating the choice between successive theories.[15] In Sections X and XII I have argued that the parties to such debates inevitably see differently certain of the experimental or observational situations to which both have recourse. Since the vocabularies in which they discuss such situations consist, however, predominantly of the same terms, they must be attaching some of those terms to nature differently, and their communication is inevitably only partial. As a result, the superiority of one theory to another is something that cannot be proved in the debate. Instead, I have insisted, each party must try, by persuasion, to convert the other. Only philosophers have seriously misconstrued the intent of these parts of my argument. A number of them, however, have reported that I believe the following:[16] the proponents of incommensurable theories cannot communicate with each other at all; as a result, in a debate over theory-choice there can be no recourse to *good* reasons; instead theory must be chosen for reasons that are ultimately personal and subjective; some sort of mystical apperception is responsible for the decision actually reached. More than any other parts of the book, the passages on which these misconstructions rest have been responsible for charges of irrationality.

Consider first my remarks on proof. The point I have been trying to make 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. In the latter, premises and rules of inference are stipulated from the start. If there is disagreement about conclusions, the parties to the ensuing debate can retrace their steps one by one, checking each against prior stipulation. At the end of that process one or the other must concede that he has made a mistake, violated a previously accepted rule. After that concession he has no recourse, and his opponent's proof is then compelling. Only if the two discover instead that they differ about the meaning or application of stipulated rules, that their prior agreement provides no sufficient basis for proof, does the debate continue in the form it inevitably takes during scientific revolutions. That debate is about premises, and its recourse is to persuasion as a prelude to the possibility of 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. If two men disagree, for example, about the relative fruitfulness of their theories, or if they agree about that but disagree about the relative importance of fruitfulness and, say, scope in reaching a choice, neither can be con-

[15] The points that follow are dealt with in more detail in Secs. v and vi of "Reflections."

[16] See the works cited in note 9, above, and also the essay by Stephen Toulmin in *Growth of Knowledge*.

victed of a mistake. Nor is either being unscientific. 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. In this sense it is the community of specialists rather than its individual members that makes the effective decision. To understand why science develops as it does, one need not unravel the details of biography and personality that lead each individual to a particular choice, though that topic has vast fascination. What one must understand, however, is the manner in which a particular set of shared values interacts with the particular experiences shared by a community of specialists to ensure that most members of the group will ultimately find one set of arguments rather than another decisive.

That process is persuasion, but it presents a deeper problem. Two men who perceive the same situation differently but nevertheless employ the same vocabulary in its discussion must be using words differently. They speak, that is, from what I have called incommensurable viewpoints. How can they even hope to talk together much less to be persuasive. Even a preliminary answer to that question demands further specification of the nature of the difficulty. I suppose that, at least in part, it takes the following form.

The practice of normal science depends on the ability, acquired from exemplars, to group objects and situations into similarity sets which are primitive in the sense that the grouping is done without an answer to the question, "Similar with respect to what?" One central aspect of any revolution is, then, that some of the similarity relations change. Objects that were grouped in the same set before are grouped in different ones afterward and vice versa. Think of the sun, moon, Mars, and earth before and after Copernicus; of free fall, pendular, and planetary motion before and after Galileo; or of salts, alloys, and a sulpuhur–iron filing mix before and after Dalton. Since most objects within even the altered sets continue to be grouped together, the names of the sets are usually preserved. Nevertheless, the transfer of a subset is ordinarily part of a critical change in the network of interrelations among them. Transferring the

metals from the set of compounds to the set of elements played an essential role in the emergence of a new theory of combustion, of acidity, and of physical and chemical combination. In short order those changes had spread through all of chemistry. Not surprisingly, therefore, when such redistributions occur, two men whose discourse had previously proceeded with apparently full understanding may suddenly find themselves responding to the same stimulus with incompatible descriptions and generalizations. Those difficulties will not be felt in all areas of even their scientific discourse, but they will arise and will then cluster most densely about the phenomena upon which the choice of theory most centrally depends.

Such problems, though they first become evident in communication, are not merely linguistic, and they cannot be resolved simply by stipulating the definitions of troublesome terms. Because the words about which difficulties cluster have been learned in part from direct application to exemplars, the participants in a communication breakdown cannot say, "I use the word 'element' (or 'mixture,' or 'planet,' or 'unconstrained motion') in ways determined by the following criteria." They cannot, that is, resort to a neutral language which both use in the same way and which is adequate to the statement of both their theories or even of both those theories' empirical consequences. Part of the difference is prior to the application of the languages in which it is nevertheless reflected.

The men who experience such communication breakdowns must, however, have some recourse. The stimuli that impinge upon them are the same. So is their general neural apparatus, however differently programmed. Furthermore, except in a small, if all-important, area of experience even their neural programming must be very nearly the same, for they share a history, except the immediate past. As a result, both their everyday and most of their scientific world and language are shared. Given that much in common, they should be able to find out a great deal about how they differ. The techniques required are not, however, either straightforward, or comfortable, or parts of the scientist's normal arsenal. Scientists rarely recognize them for quite what they are, and they seldom use them for longer than is required to induce conversion or convince themselves that it will not be obtained.

Briefly put, what the participants in a communication breakdown can do is recognize each other as members of different language communities and then become translators.[17] Taking the differences between their own intra- and inter-group discourse as itself a subject for study, they can first attempt to discover the terms and locutions that, used unproblematically within each community, are nevertheless foci of trouble for inter-group discussions. (Locutions that present no such difficulties may be homophonically translated.) Having isolated such areas of difficulty in scientific communication, they can next resort to their shared everyday vocabularies in an effort further to elucidate their troubles. Each may, that is, try to discover what the other would see and say when presented with a stimulus to which his own verbal response would be different. If they can sufficiently refrain from explaining anomalous behavior as the consequence of mere error or madness, they may in time become very good predictors of each other's behavior. Each will have learned to translate the other's theory and its consequences into his own language and simultaneously to describe in his language the world to which that theory applies. That is what the historian of science regularly does (or should) when dealing with out-of-date scientific theories.

Since translation, if pursued, allows the participants in a communication breakdown to experience vicariously something of the merits and defects of each other's points of view, it is a potent tool both for persuasion and for conversion. But even persuasion need not succeed, and, if it does, it need not be

[17] The already classic source for most of the relevant aspects of translation is W. V. O. Quine, *Word and Object* (Cambridge, Mass., and New York, 1960), chaps. i and ii. But Quine seems to assume that two men receiving the same stimulus must have the same sensation and therefore has little to say about the extent to which a translator must be able to *describe* the world to which the language being translated applies. For the latter point see, E. A. Nida, "Linguistics and Ethnology in Translation Problems," in Del Hymes (ed.), *Language and Culture in Society* (New York, 1964), pp. 90–97.

accompanied or followed by conversion. The two experiences are not the same, an important distinction that I have only recently fully recognized.

To persuade someone is, I take it, to convince him that one's own view is superior and ought therefore supplant his own. That much is occasionally achieved without recourse to anything like translation. In its absence many of the explanations and problem-statements endorsed by the members of one scientific group will be opaque to the other. But each language community can usually produce from the start a few concrete research results that, though describable in sentences understood in the same way by both groups, cannot yet be accounted for by the other community in its own terms. If the new viewpoint endures for a time and continues to be fruitful, the research results verbalizable in this way are likely to grow in number. For some men such results alone will be decisive. They can say: I don't know how the proponents of the new view succeed, but I must learn; whatever they are doing, it is clearly right. That reaction comes particularly easily to men just entering the profession, for they have not yet acquired the special vocabularies and commitments of either group.

Arguments statable in the vocabulary that both groups use in the same way are not, however, usually decisive, at least not until a very late stage in the evolution of the opposing views. Among those already admitted to the profession, few will be persuaded without some recourse to the more extended comparisons permitted by translation. Though the price is often sentences of great length and complexity (think of the Proust-Berthollet controversy conducted without recourse to the term 'element'), many additional research results can be *translated* from one community's language into the other's. As translation proceeds, furthermore, some members of each community may also begin vicariously to understand how a statement previously opaque could seem an explanation to members of the opposing group. The availability of techniques like these does not, of course, guarantee persuasion. For most people translation is a threatening process, and it is entirely foreign to normal science.

Counter-arguments are, in any case, always available, and no rules prescribe how the balance must be struck. Nevertheless, as argument piles on argument and as challenge after challenge is successfully met, only blind stubbornness can at the end account for continued resistance.

That being the case, a second aspect of translation, long familiar to both historians and linguists, becomes crucially important. To translate a theory or worldview into one's own language is 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. That transition is not, however, one that an individual may make or refrain from making by deliberation and choice, however good his reasons for wishing to do so. Instead, at some point in the process of learning to translate, he finds that the transition has occurred, that he has slipped into the new language without a decision having been made. Or else, like many of those who first encountered, say, relativity or quantum mechanics in their middle years, he finds himself fully persuaded of the new view but nevertheless unable to internalize it and be at home in the world it helps to shape. Intellectually such a man has made his choice, but the conversion required if it is to be effective eludes him. He may use the new theory nonetheless, but he will do so as a foreigner in a foreign environment, an alternative available to him only because there are natives already there. His work is parasitic on theirs, for he lacks the constellation of mental sets which future members of the community will acquire through education.

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.

6. *Revolutions and Relativism*

One consequence of the position just outlined has particularly bothered a number of my critics.[18] They find my viewpoint relativistic, particularly as it is developed in the last section of this book. My remarks about translation highlight the reasons for the charge. The proponents of different theories are like the members of different language-culture communities. Recognizing the parallelism suggests that in some sense both groups may be right. Applied to culture and its development that position is relativistic.

But applied to science it may not be, and it is in any case far from *mere* relativism in a respect that its critics have failed to see. Taken as a group or in groups, practitioners of the developed sciences are, I have argued, fundamentally puzzle-solvers. Though the values that they deploy at times of theory-choice derive from other aspects of their work as well, the demonstrated ability to set up and to solve puzzles presented by nature is, in case of value conflict, the dominant criterion for most members of a scientific group. Like any other value, puzzle-solving ability proves equivocal in application. Two men who share it may nevertheless differ in the judgments they draw from its use. But the behavior of a community which makes it preeminent will be very different from that of one which does not. In the sciences, I believe, the high value accorded to puzzle-solving ability has the following consequences.

Imagine an evolutionary tree representing the development of the modern scientific specialties from their common origins in, say, primitive natural philosophy and the crafts. A line drawn up that tree, never doubling back, from the trunk to the tip of some branch would trace a succession of theories related by descent. Considering any two such theories, chosen from points not too near their origin, it should be easy to design a list of criteria that would enable an uncommitted observer to distinguish the earlier from the more recent theory time after time. Among

[18] Shapere, "Structure of Scientific Revolutions," and Popper in *Growth of Knowledge*.

the most useful would be: accuracy of prediction, particularly of quantitative prediction; the balance between esoteric and everyday subject matter; and the number of different problems solved. Less useful for this purpose, though also important determinants of scientific life, would be such values as simplicity, scope, and compatibility with other specialties. Those lists are not yet the ones required, but I have no doubt that they can be completed. If they can, then scientific development is, like biological, a unidirectional and irreversible process. Later scientific theories are better than earlier ones for solving puzzles in the often quite different environments to which they are applied. That is not a relativist's position, and it displays the sense in which I am a convinced believer in scientific progress.

Compared with the notion of progress most prevalent among both philosophers of science and laymen, however, this position lacks an essential element. A scientific theory is usually felt to be better than its predecessors not only in the sense that it is a better instrument for discovering and solving puzzles but also because it is somehow a better representation of what nature is really like. One often hears that successive theories grow ever closer to, or approximate more and more closely to, the truth. Apparently generalizations like that refer not to the puzzle-solutions and the concrete predictions derived from a theory but rather to its ontology, to the match, that is, between the entities with which the theory populates nature and what is "really there."

Perhaps there is some other way of salvaging the notion of 'truth' for application to whole theories, but this one will not do. There is, I think, no theory-independent way to reconstruct phrases like 'really there'; the notion of a match between the ontology of a theory and its "real" counterpart in nature now seems to me illusive in principle. Besides, as a historian, I am impressed with the implausibility of the view. I do not doubt, for example, that Newton's mechanics improves on Aristotle's and that Einstein's improves on Newton's as instruments for puzzle-solving. But I can see in their succession no coherent direction of ontological development. On the contrary, in some

important respects, though by no means in all, Einstein's general theory of relativity is closer to Aristotle's than either of them is to Newton's. Though the temptation to describe that position as relativistic is understandable, the description seems to me wrong. Conversely, if the position be relativism, I cannot see that the relativist loses anything needed to account for the nature and development of the sciences.

7. *The Nature of Science*

I conclude with a brief discussion of two recurrent reactions to my original text, the first critical, the second favorable, and neither, I think, quite right. Though the two relate neither to what has been said so far nor to each other, both have been sufficiently prevalent to demand at least some response.

A few readers of my original text have noticed that I repeatedly pass back and forth between the descriptive and the normative modes, a transition particularly marked in occasional passages that open with, "But that is not what scientists do," and close by claiming that scientists ought not do so. Some critics claim that I am confusing description with prescription, violating the time-honored philosophical theorem: 'Is' cannot imply 'ought.'[19]

That theorem has, in practice, become a tag, and it is no longer everywhere honored. A number of contemporary philosophers have discovered important contexts in which the normative and the descriptive are inextricably mixed.[20] 'Is' and 'ought' are by no means always so separate as they have seemed. But no recourse to the subtleties of contemporary linguistic philosophy is needed to unravel what has seemed confused about this aspect of my position. The preceding pages present a viewpoint or theory about the nature of science, and, like other philosophies of science, the theory has consequences for the way in which scientists should behave if their enterprise is to succeed. Though

[19] For one of many examples, see P. K. Feyerabend's essay in *Growth of Knowledge*.

[20] Stanley Cavell, *Must We Mean What We Say?* (New York, 1969), chap. i.

it need not be right, any more than any other theory, it provides a legitimate basis for reiterated 'oughts' and 'shoulds.' Conversely, one set of reasons for taking the theory seriously is that scientists, whose methods have been developed and selected for their success, do in fact behave as the theory says they should. My descriptive generalizations are evidence for the theory precisely because they can also be derived from it, whereas on other views of the nature of science they constitute anomalous behavior.

The circularity of that argument is not, I think, vicious. The consequences of the viewpoint being discussed are not exhausted by the observations upon which it rested at the start. Even before this book was first published, I had found parts of the theory it presents a useful tool for the exploration of scientific behavior and development. Comparison of this postscript with the pages of the original may suggest that it has continued to play that role. No merely circular point of view can provide such guidance.

To one last reaction to this book, my answer must be of a different sort. A number of those who have taken pleasure from it have done so less because it illuminates science than because they read its main theses as applicable to many other fields as well. I see what they mean and would not like to discourage their attempts to extend the position, but their reaction has nevertheless puzzled me. To the extent that the book portrays scientific development as a succession of tradition-bound periods punctuated by non-cumulative breaks, its theses are undoubtedly of wide applicability. But they should be, for they are borrowed from other fields. Historians of literature, of music, of the arts, of political development, and of many other human activities have long described their subjects in the same way. Periodization in terms of revolutionary breaks in style, taste, and institutional structure have been among their standard tools. If I have been original with respect to concepts like these, it has mainly been by applying them to the sciences, fields which had been widely thought to develop in a different way. Conceivably the notion of a paradigm as a concrete achievement, an exemplar, is a second contribution. I suspect, for example, that some of the notorious difficulties surrounding the notion of style in the

arts may vanish if paintings can be seen to be modeled on one another rather than produced in conformity to some abstracted canons of style.[21]

This book, however, was intended also to make another sort of point, one that has been less clearly visible to many of its readers. Though scientific development may resemble that in other fields more closely than has often been supposed, it is also strikingly different. To say, for example, that the sciences, at least after a certain point in their development, progress in a way that other fields do not, cannot have been all wrong, whatever progress itself may be. One of the objects of the book was to examine such differences and begin accounting for them.

Consider, for example, the reiterated emphasis, above, on the absence or, as I should now say, on the relative scarcity of competing schools in the developed sciences. Or remember my remarks about the extent to which the members of a given scientific community provide the only audience and the only judges of that community's work. Or think again about the special nature of scientific education, about puzzle-solving as a goal, and about the value system which the scientific group deploys in periods of crisis and decision. The book isolates other features of the same sort, none necessarily unique to science but in conjunction setting the activity apart.

About all these features of science there is a great deal more to be learned. Having opened this postscript by emphasizing the need to study the community structure of science, I shall close by underscoring the need for similar and, above all, for comparative study of the corresponding communities in other fields. How does one elect and how is one elected to membership in a particular community, scientific or not? What is the process and what are the stages of socialization to the group? What does the group collectively see as its goals; what deviations, individual or collective, will it tolerate; and how does it control the impermissible aberration? A fuller understanding of science will de-

[21] For this point as well as a more extended discussion of what is special about the sciences, see T. S. Kuhn, "Comment [on the Relations of Science and Art]," *Comparative Studies in Philosophy and History*, XI (1969), 403–12.

pend on answers to other sorts of questions as well, but there is no area in which more work is so badly needed. Scientific knowledge, like language, is intrinsically the common property of a group or else nothing at all. To understand it we shall need to know the special characteristics of the groups that create and use it.

Index

This index has been prepared by Peter J. Riggs, and both author and publisher are indebted to him for recommending this addition and seeing it into print.