Second Thoughts on Paradigms

Thomas Kuhn

「典範論」的再思考

孔恩

It has now been several years since a book of mine, The Structure of Scientific Revolutions, was published. Reactions to it have been varied and occasionally strident, but the book continues to be widely read and much discussed. By and large I take great satisfaction from the interest it has aroused, including much of the criticism. One aspect of the response does, however, from time to time dismay me. Monitoring conversations, particularly among the book's enthusiasts, I have sometimes found it hard to believe that all parties to the discussion had been engaged with the same volume. Part of the reason for its success is, I regretfully conclude, that it can be too nearly all things to all people.o much responsible as its introduction of the term "paradigm" a word that figures more than an other, excepting the grammaticalFor that excessive plasticity, no aspect of the book is so much responsible as its introduction of the term "paradigm" a word that figures more than an other, excepting the grammatical particles, in its pages. Challenged to explain the absence of an index, I regularly point out that its most frequently consulted entry would be: "paradigm, 1-172, passim." Critics, whether sympathetic or not, have been unanimous in underscoring the large number of different senses in which the term is used. One commentator, who thought the matter worth systematic scrutiny, prepared a partial subject index and found at least twenty-two different usages, ranging from "a concrete scientific achievement" (p.11) to a "characteristic set of beliefs and preconceptions" (p. 17), the latter including instrumental, theoretical, and metaphysical commitments together (pp. 39-42). Though neither the compiler of that index nor I think the situation so desperate as those divergences suggest, clarification is obviously called for. Nor will clarification by itself suffice. Whatever their number, the usages of "paradigm" in the book divide, into two sets which require both different names and separate discussion. Our sense of "paradigm" is global, embracing all the shared commitments of a scientific group; the other isolates a particularly important sort of commitment and is thus a subset of the first. In what follows I shall try initially to disentangle them and then to scrutinize the one that I believe most urgently needs philosophical attention. However imperfectly I understood paradigms when I wrote the book, I still think them worth much attention.

In the book the term "paradigm" enters in close proximity, both physical and logical, to the phrase "scientific community" (pp.10-11). A paradigm is what the members of a scientific community, and they alone, share. Conversely, it is their possession of a common paradigm that constitutes a scientific community of a group of otherwise disparate men. As empirical generalizations, both those statements can be defended. But in the book they function at least partly as definitions, and the result is a circularity with at least a few vicious consequences. If the term "paradigm" is to be successfully explicated, scientific communities must first be recognized as having an independent existence.

In fact, the identification and study of scientific communities has recently emerged as a significant research subject among sociologists. Preliminary results, many of them still unpublished, suggest that the requisite empirical techniques are nontrivial, but some are already in hand, and others are sure to be developed. Most practicing scientists respond at once to questions about their community affiliations, taking it for granted that responsibility for the various current specialties and research techniques is distributed among groups of at least roughly determinate membership. I shall therefore assume that more systematic means for their identification will be forthcoming and content myself here with a brief articulation of an intuitive notion of community, one widely shared be scientists, sociologists, and a number of historians of science.

A scientific community consists, in this view, of the practitioners of a scientific specialty. Bound together by common elements in their education and apprenticeship, they see themselves and are seen by others as the men responsible for the pursuit of a set of shared goals, including the training of their successors. Such communities are characterized by the relative fullness of communication within the group and by the relative unanimity of the group's judgment in professional matters. To a remarkable extent the members of a given community will have absorbed the same literature and drawn similar lessons from it. Because the attention of different communities is focused on different matters, professional communication across group lines is likely to be arduous, often gives rise to misunderstanding, and may, if pursued, isolate significant disagreement.

Clearly, communities in this sense exist at numerous levels. Perhaps all natural scientists form a community. (We ought not, I think, allow the storm surrounding C. P. Snow to obscure those points about which he has said the obvious.) At an only slightly lower level, the main scientific professional groups provide examples of communities: physicists, chemists, astronomers, zoologists, and the like. For these major communities group 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 the 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 difficulties emerge. How, prior to its public acclaim, would an outsider have isolated the phage group? For this, one must have recourse to attendance at summer institutes and special conferences, to preprint distribution lists, and above all to formal and informal communication networks, including the linkages 'among citations'. I take it that the job can and will be done, and that it will typically yield communities of perhaps a hundred members, sometimes significantly fewer. Individual scientists, particularly the ablest, will belong to several such groups, either simultaneously or in succession. Though it is not yet clear just how far empirical analysis cantake us, there is excellent reason to suppose that the scientific enterprise is distributed among and carried forward by communities of this sort.

Let me now suppose that we have, by whatever techniques, identified one such community. What shared elements account for the relatively unproblematic character of professional communication and for the relative unanimity of professional judgment? To this question The Structure of Scientific Revolutions licences the answer "a paradigm" or "a set of paradigms." That is one of the two main senses in which the term occurs in the book. For it I might now adopt the notation "paradigm@" but less confusion will result if I instead replace it with the phrase "disciplinary matrix"--"disciplinary" because it is the common possession of the practitioners of a professional discipline and "matrix" because it is composed of ordered elements of various sorts, each requiring further specification. Constituents of the disciplinary matrix include most or all of the objects of group commitment described in the book as paradigms, parts of paradigms, or paradigmatic. I shall not at this time even attempt an exhaustive list but will instead briefly identify three of these which, because they are central to the cognitive operation of the group, should particularly concern philosophers of science. Let me refer to them as symbolic generalizations, models, and exemplars.

The first two are already familiar objects of philosophical attention. Symbolic generalizations, in particular, are those expressions, deployed without question by the group, which can readily be cast in some logical form like (x)(y)(z)Φ(x, y, z). They are the formal, or the readily formalizable, components of the disciplinary matrix. Models, about which I shall have nothing further to say in this paper, are what provide the group with preferred analogies or, when deeply held, with an ontology. At one extreme they are heuristic: the electric circuit may fruitfully be regarded as a steady state hydrodynamic system, or a gas behaves like a collection of microscopic billiard balls in random motion. At the other, they are the objects of metaphysical commitment: the heat of a body is the kinetic energy of its constituent particles, or, more obviously metaphysical, all perceptible phenomena are due to the motion and interaction of qualitatively neutral atoms in the void." Exemplars, finally, are concrete problem solutions, accepted by the group as, in a quite usual sense, paradigmatic. Many of you will already have guessed that the term "exemplar" provides a new name for thesecond, and more fundamental, sense of "paradigm" in the book.

To understand how a scientific community functions as a producer and validator of sound knowledge, we must ultimately, I think, understand the operation of at least these three components of the disciplinary matrix. Alterations in any one can result in changes of scientific behavior affecting both the locus of a group's research and its standards of verification. Here I shall not attempt to defend a thesis quite so general. My primary concern is now with exemplars. To make room for them, however, I must first say something about symbolic generalizations.

In the sciences, particularly in physics, generalizations are often found already in symbolic form: f = ma, I = V/R. Others-are ordinarily expressed in words: "action equals reaction," "chemical composition is in fixed proportions by weight," or "all cells come from cells." No one will question that the members of a scientific community do routinely deploy expressions like these in their work, that they ordinarily do so without felt need for special justification, and that they are seldom challenged at such points by other members of their group. That behavior is important, for without a shared commitment to a set of symbolic generalizations, logic and mathematics could not routinely be applied in the community's work. The example of .taxonomy suggests that a science can exist with few, perhaps with no, such generalizations. I shall later suggest how this could be the case. But I see no reason to doubt the widespread impression that the power of a science increases with the number of symbolic generalizations its practitioners have at their disposal.

Note, however, how small a measure of agreement we have yet attributed to the members of our community. When I say they share a commitment to, say, the symbolic generalization f = ma, I mean only that they will raise no difficulties for the man who inscribes the four symbols f, =, m, and a in succession on a line, who manipulates the resulting expression by logic and mathematics, and who exhibits a still symbolic result. For us at this point in the discussion, though not for the scientists who use them, these symbols and the expressions formed by compounding them are uninterpreted, still empty of empirical meaning or application. A shared commitment to a set of generalizations justifies logical and mathematical manipulation and introduces commitment to the result. It need not, however, imply agreement about the manner in which the symbols, individually and collectively, are to be correlated with the results of experiment and observation. To this extent the shared symbolic generalizations function as yet like expressions in a pure mathematical system.

The analogy between a scientific theory and a pure mathematical system has been widely exploited in twentieth-century philosophy of science and has been responsible for some extremely interesting results. But it is only an analogy and can therefore be misleading. I believe that in several respects we have been victimized by it. One of them has immediate relevance to my argument.

When an expression like f = ma appears in a pure mathematical system, it is, so to speak, there once and for all. If, that is, it enters into the solution of a mathematical problem posed within the system, it always enters in the form f = ma or in a form reducible to that one by the substitutivity of identities or by some other syntactic substitution rule. In the sciences symbolic generalizations ordinarily behave very differently. They are not so much generalizations as generalization-sketches, schematic forms whose detailed symbolic expression varies from one application to the next. For the problem of free fall, f = ma becomes mg = mds/dt. For the simple pendulum, it becomes mgsin = - mds/dt. For coupled harmonic oscillators it becomes two equations, the first of which may be written mds/dt+ks=k(d+s-s). More interesting mechanical problems, for example, the motion of gyroscope, would display still greater disparity between f = ma and the actual symbolic generalization to which logic and mathematics are applied; but the point should already be clear. Though uninterpreted symbolic expressions are the common possession of the members of a scientific community, and though it is such expressions which provide the group with an entry point for logic and mathematics, it is not to the shared generalization that these tools are applied but to one or another special version of it. In a sense, each such class requires a new formalism.

An interesting conclusion follows, one with likely relevance to the status of theoretical terms. Those philosophers who exhibit scientific theories as uninterpreted formal systems often remark that empirical reference enters such theories from the bottom up, moving from an empirically meaningful basic vocabulary into the theoretical terms. Despite the well-known difficulties that cluster about the notion of a basic vocabulary, I cannot doubt the importance of that route in the transformation of an uninterpreted symbol into the sign for a particular physical concept. But it is not the only route. Formalisms in science- also attach to nature at the top, without intervening deduction which eliminates theoretical terms. Before he can begin the logical and mathematical manipulations which eventuate with the prediction of meter readings, the scientist must inscribe the particular form of / = ma that applies to, say, the vibrating string or the particular form of the Schrodinger equation which applies to, say, the helium atom in a magnetic field. Whatever procedure he employs in doing so, it cannot be purely syntactic. Empirical content must enter formalized theories from the top as well as the bottom.

One cannot, I think, escape this conclusion by suggesting that the Schr5dinger equation or f = ma be construed as an abbreviation for a conjunction of the numerous particular symbolic forms which these expressions take for application to particular physical problems. In the first place, scientists would still require criteria to tell them which particular symbolic version should be applied to which problem, and these criteria, like the correlation rules that are said to transport meaning from a basic vocabulary to theoretical terms, would be a vehicle for empirical content. Besides, no conjunction of particular symbolic forms would exhaust what the members of a scientific community can properly be said to know about how to apply symbolic generalizations. Confronted with a new problem, they can often agree on the particular symbolic expression appropriate to it, even though none of them has seen that particular expression before.

Any account of the cognitive apparatus of a scientific community may reasonably be asked to tell us something about the way in which the group's members, in advance of directly relevant empirical evidence, identify the special formalism appropriate to a particular problem, especially to a new problem. That clearly is one of the functions which scientific knowledge does serve. It does not, of course, always do so correctly; there is room, indeed need, for empirical checks on a special formalism proposed for a new problem. The deductive steps and the comparison of their end products with experiment remain prerequisites of science. But special formalisms are regularly accepted as plausible or rejected as implausible in advance of experiment. With remarkable frequency, furthermore, the community's judgments prove to be correct. Designing a special formalism, a new version of the formalization, cannot therefore be quite like inventing a new theory. Among other things, the former can be taught as theory invention cannot. That is what the problems at the ends of chapters in science texts are principally for. What can it be that students learn while solving them?

To that question most of the remainder of this paper is devoted, but I shall approach it indirectly, asking at first a more usual one: How do scientists attach symbolic expressions to nature? That is, in fact, two questions in one, for it may be asked either about a special symbolic generalization designed for a particular experimental situation or about a singular symbolic consequence of that generalization deduced for comparison with experiment. For present purposes, however, we may treat these two questions as one. In scientific practice, also, they are ordinarily answered together.

Since the abandonment of hope for a sense-datum language, the usual answer to this question has been in terms if correspondence rules. These have ordinarily been taken to be either operational definitions of scientific terms or else a set of necessary and sufficient conditions for the terms' applicability. I do not myself doubt that the examination of a given scientific community would disclose a number of such rules shared by its member Probably a few others could legitimately be included from close observation of their behavior. But, for reasons I have given elsewhere and shall advert behavior below, I do doubt that the correspondence rules discovered in this way would be nearly sufficient in number or force to account for the actual correlations between formalism and experiment made regularly and unproblematically by members of the group. If the philosopher wants an adequate body of correspondence rules, he will have to supply most of them for himself.

Almost surely that is a job he can do. Examining the collected examples of past community practice, the philosopher may reason ably expect to construct a set of correspondence rules adequate, in conjunction with known symbolic generalizations, to account for them all. Very likely he would be able to construct several alternates sets. Nevertheless, he ought to be extraordinarily wary about describing any one of them as a reconstruction of the rules held by the community under study. Though each of his sets of rules would be equivalent with respect to the community's past practice, they need not be equivalent when applied to the very next problem faced by the discipline. In that sense they would be reconstructions of slightly different theories, none of which need be the one held by the group. The philosopher might well, by behaving as a scientist, have improved the group's theory, but he would not, as a philosopher, have analyzed it.

Suppose, for example, that the philosopher is concerned with Ohm's law, $I = V/R$, and that he knows that the members of the group he studies measure voltage with an electrometer and current with a galvanometer. Seeking a correspondence rule for resistance, he may choose the quotient of voltage divided by current, in which case Ohm's law becomes a tautology. Or he may instead choose to correlate the value of resistance with the results of measurements made on a Wheatstone Bridge, in which case Ohm's law provides information about nature. For past practice the two reconstructions may be equivalent, but they will not dictate the same future behavior. Imagine, in particular, that an especially adept experimentalist in the community applies higher voltages than any realized before and discovers that the voltage-to-current ratio changes gradually at high voltage. According to the second, the Wheatstone Bridge, reconstruction, he has discovered that there are deviations from Ohm's law at high voltage. On the first reconstruction, however, Ohm's law is a tautology and deviations from it are unimaginable. The experimentalist has discovered, not a deviation from the law, but rather that resistance changes with voltage. The two reconstructions lead to different localizations of the difficulty and to different patterns of follow-up research.

Nothing in the preceding discussion proves that there is no set of correspondence rules adequate to explain the behavior of the community under study. A negative of that sort scarcely can be proven. But the discussion may lead us to take a bit more seriously some aspects of scientific training and behavior that philosophers have often managed to look right through. Very few correspondence rules are to be found in science texts or science teaching. How can the members of a scientific community have acquired a sufficient set? It is also noteworthy that if asked by a philosopher to provide such rules, scientists regularly deny their relevance and thereafter sometimes grow uncommonly inarticulate. When they cooperate at all, the rules they produce may vary from one member of the community to another, and all may be defective.*0ne begins to wonder whether more than a few such rules are deployed in community practice, whether there is not some alternate way in which scientists correlate their symbolic expressions with nature.

A phenomenon familiar both to students of science and to historians of science provides a clue. Having been both, I shall speak from experience. Students of physics regularly report that they have read through a chapter of their text, understood it perfectly, but nonetheless had difficulty solving the problems at the end of the chapter. Almost invariably their difficulty is in setting up the appropriate equations, in relating the words and examples given in the text to the particular problems they are asked to solve. Ordinarily, also, those difficulties dissolve in the same way. The student discovers a way to see his problem as like a problem he has already encountered. Once that likeness or analogy has been seen, only manipulative difficulties remain.

The same pattern shows clearly in the history of science. Scientists model one problem solution on another, often with only a 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, like that of Galileo's pendulum, would rise only to the height from which the center of gravity of the extended pendulum had begun to fall. Finally, Daniel Bernoulli, still with no aid from Newton's laws, discovered how to make the flow of water from an orifice in a storage tank resemble Huyghens's 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 afterwards moves separately upward to the maximum height obtainable with the velocity it possessed at the end of the interval of descent. The ascent of the center of gravity of the separate 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.

Lacking time to multiply examples, I suggest that an acquired ability to see resemblances between apparently disparate problems plays in the sciences a significant part of the role usually attributed to correspondence rules. Once a new problem is seen to be analogous to a problem previously solved, both an appropriate formalism and a new way of attaching its symbolic consequences to nature follow. Having seen the resemblance, one simply uses the attachments that have proved effective before. That ability to recognize group-licensed resemblances is, I think, the main thing students acquire by doing problems, whether with pencil and paper or in a well-designed laboratory. In the course of their training a vast number of such exercises are set for them, and students entering the same specialty regularly do very nearly the same ones, for example, the inclined plane, the conical pendulum, Kepler ellipses, and so on. These concrete problems with their solutions are what I previously referred to as exemplars, a community's standard examples. They constitute the third main sort of cognitive component of the disciplinary matrix, and they illustrate the second main function of the term "paradigm" in The Structure of Scientific Revolutions. Acquiring an arsenal of exemplars, just as much as learning symbolic generalizations, is integral to the process by which a student gains access to the cognitive achievements of his disciplinary group. Without exemplars he would never learn much of what the group knows about such fundamental concepts as force and field, element and compound, or nucleus and cell.

I shall shortly attempt, by means of a simple example, to explicate the notion of a learned similarity relationship, an acquired perception of analogy. Let me first, however, sharpen the problem at which that explication will be aimed. It is a truism that anything is similar to, and also different from, anything else. It depends, we usually say, on the criteria. To the man who speaks of similarity or of analogy, we therefore at once pose the question: similar with respect to what? In this case, however, that is just the question that must not be asked, for an answer would at once provide us with correspondence rules. Acquiring exemplars would teach the student nothing that such rules, in the form of criteria of resemblance, could not equally well have supplied. Doing problems would then be mere practice in the application of rules, and there would be no need for talk of similarity.

Doing problems, however, I have already argued, is not like that. Much more nearly it resembles the child's puzzle in which one is asked to find the animal shapes or faces hidden in the drawing of shrubbery or clouds. The child seeks forms that are like those of the animals or faces he knows. Once they are found, they do not again retreat into the background, for the child's way of seeing the picture has been changed. In the same way, the science student, confronted with a problem, seeks to see it as like one or more of the exemplary problems he has encountered before. Where rules exist to guide him, he, of course, deploys them. But his basic criterion is a perception of similarity that is both logically and psycho logically prior to any of the numerous criteria by which that same identification of similarity might have been made. After the similarity has been seen, one may ask for criteria, and it is then often worth doing so. But one need not. The mental or visual set acquired while learning to see two problems as similar can be applied directly. Under appropriate circumstances, I now want to argue, there is a means of processing data into similarity sets which does not depend on a prior answer to the question, similar with respect to what?

My argument begins with a brief digression on the term "data." Philologically it derives from "the given." philosophically, for reasons deeply engrained in the history of .epistemology, it isolates the minimal stable elements provided by our senses. Though we no longer hope for a sense-datum language, phrases like "green there," "triangle here," or "hot down there" continue to connote our paradigms for a datum, the given in experience. In several respects, they should play this role. We have no access to elements of experience more minimal than these. Whenever we consciously process data, whether to identify an object, to discover a law, or to invent a theory, we necessarily manipulate sensations of this sort or compounds of them. Nevertheless, from another point of view, sensations and their elements are not the given. Viewed theoretically rather than experientially, that title belongs rather to stimuli. Though we have access to them only indirectly, via scientific theory, it is stimuli, not sensations, that impinge on us as organisms. A vast amount of neural processing takes place between our receipt of a stimulus and the sensory response which is our datum. None of this would be worth saying if Descartes had been night in positing a one-to-one correspondence between stimuli and sensations. But we know that nothing of the sort exists. The perception of a given color can be evoked by an infinite number of differently combined wavelengths. Conversely, a given stimulus can evoke a variety of sensations, the image of a duck in one recipient, the image of a rabbit in another. Nor are responses like these entirely innate. One can learn to discriminate colors or patterns which were indistinguishable prior to training. To an extent still unknown, the production of data from stimuli is a learned procedure. After the learning process, the same stimulus evokes a different datum. I conclude that, though data are the minimal elements of our individual experience/they need be shared responses to a given stimulus only within the membership of a relatively homogeneous community, educational, scientific, or linguistic.

Return now to my main argument, but not to scientific examples. Inevitably the latter prove excessively complex. Instead I ask that you imagine a small child on a walk with his father in a zoological garden. The child has previously learned to recognize birds and to discriminate robin redbreasts. During the "afternoon now at hand, he will learn for the first time to identify swans, geese, and ducks. Anyone who has taught a child under such circumstances knows that the primary pedagogic tool is ostension. Phrases like "all swans are white" may play a role, but they need not. I shall for the moment omit them from consideration, my object being to isolate a different mode of learning in its purest form. Johnny's education then proceeds as follows. Father points to a bird, saying, "Look, Johnny, there's a swan." A short time later Johnny himself points to a bird, saying, "Daddy, another swan." He has not yet, however, learned what swans are and must be corrected: "No, Johnny, that's a goose." Johnny's next identification of a swan proves to be correct, but his next "goose" is, in fact, a duck, and he is again set straight. After a few more such encounters, however, each with its appropriate correction or reinforcement, Johnny's ability to identify these waterfowl is as great, as his father's. Instruction has been quickly completed.

I ask now what has happened to Johnny, and I urge the plausibility of the following answer. During the afternoon, part of the neural mechanism by which he processes visual stimuli has been reprogrammed, and the data he receives from stimuli which would all earlier have evoked "bird" have changed. When he began his walk, the neural program highlighted the differences between individual swans as much as those between swans and geese. By the end of the walk, features like the length and curvature of the swan's neck have been highlighted and others have been suppressed so that swan data match each other and differ from goose and duck data as they had not before. Birds that had previously all looked alike (and also different) are now grouped in discrete clusters in perceptual space.

A process of this sort can readily be modeled on a computer; I am in the early stages of such an experiment myself. A stimulus, in the form of a string of n ordered digits, is fed to the machine. There it is transformed to a datum by the application of a preselected transformation to each of the n digits, a different transformation being applied to each position in the string. Every datum thus obtained is a string of n numbers, a position in what I shall call an n-dimensional quality space. In this space the distance between two data, measured with a euclidean or a suitable noneuclidean metric, represents their similarity. Which stimuli transform to similar or nearby data depends, of course, on the choice of transformation functions. Different sets of functions produce different clusters of data, different patterns of similarity and difference, in perceptual space. But the transformation functions need not be manmade. If the machine is given stimuli which can be grouped in clusters and if it is informed which stimuli must be placed in the same and which in different clusters, it can design an appropriate set of transformation functions for itself. Note that both conditions are essential. Not all stimuli can be transformed to form data clusters. Even when they can, the machine, like the child, must be told at first which ones belong together and which apart. Johnny did not discover for himself that there were swans,

Figure 1

Figure 2

geese, and ducks. Rather he was taught it.

If we now represent Johnny's perceptual space in a two-dimensional diagram, the process he has undergone is rather like the transition from figure I to figure 2. In the first, ducks, geese, and swans are mixed together. In the second, they have clustered in discrete sets with appreciable distances between them. Since Johnny's father has, in effect, told him that ducks, geese, and swans are members of discrete natural families, Johnny has every right to expect that all future ducks, geese, and swans will fall naturally into or at the edge of one of these families, and that he will encounter no datum that falls in the region midway between them. That expectation may be violated, perhaps during a visit to Australia. But it will serve him well while he remains a member of the community that has discovered from experience the utility and viability of these particular perceptual discriminations and has transmitted the ability to make them from one generation to the next.

By being programmed to recognize what his prospective community already knows, Johnny has acquired consequential information. He has learned that geese, ducks, and swans form discrete natural families and that nature offers no swan-geese or goose ducks. Some quality constellations go together; others are not found at all. If the qualities in his clusters include aggressiveness, his afternoon in the park may have had behavioral as well as every day zoological functions. Geese, unlike swans and ducks, hiss and bite. What Johnny has learned is thus worth knowing. But does he know what the terms ''goose,'' "duck," and "swan" mean? In any useful sense, yes, for he can apply these labels unequivocally and without effort, drawing behavioral conclusions from their application, either directly, or via general statements. On the other hand, he has learned all this without acquiring, or at least without needing to acquire, even one criterion for identifying swans, geese, or ducks. He can point to a swan and tell you there must be water nearby, but he may well be unable to tell you what a swan is.

Johnny, in short, has learned to apply symbolic labels to nature without anything like definitions or correspondence rules. In their absence he employs a learned but nonetheless primitive perception of similarity and difference. While acquiring the perception, he has learned something about nature. This knowledge can thereafter be embedded, not in generalizations or rules, but in the similarity relationship itself. I do not, let me emphasize, at all suppose Johnny's technique is the only one by which knowledge is acquired and stored. Nor do I think it likely that very much human knowledge is acquired and stored with so little recourse to verbal generalizations. But I do urge the recognition of the integrity of acognitive process like the one just outlined. In combination with more familiar processes, like symbolic generalization and modeling, it is, I think, essential to an adequate reconstruction of scientific knowledge.

Need I now say that the swans, geese, and ducks which Johnny encountered during his walk with father were what I have been calling exemplars? Presented to Johnny with their labels attached, they were solutions to a problem that the members of his prospective community had already resolved. Assimilating them is part of the socialization procedure by which Johnny is made part of that community and, in the process, learns about the world which the community inhabits. Johnny is, of course, no scientist, nor is what he has learned yet science. But he may well become a scientist, and the technique employed on his walk will still be viable. That he does, in fact, use it will be most obvious if he becomes a taxonomist. The herbaria, without which no botanist can function, are storehouses for professional exemplars, and their history is coextensive with that of the discipline they support. But the same technique, if in a less pure form, is essential to the more abstract sciences as well. I have already argued that assimilating solutions to such problems as the inclined plane and the conical pendulum is part of learning what Newtonian physics is. Only after a number of such problems have been assimilated, can a student or a professional proceed to identify other Newtonian problems for himself. That assimilation of examples is, furthermore, part of what enables him to isolate the forces, masses, and constraints within a new problem and to write down a formalism suitable for its solution. Despite its excessive simplicity, Johnny's case should suggest why I continue to insist that shared examples have essential cognitive functions prior to a specification of criteria with respect to which they are exemplary.

I conclude my argument by returning to a crucial question discussed earlier in connection with symbolic generalizations. Suppose scientists to assimilate and store knowledge in shared examples, need the philosopher concern himself with the process? May derive correspondence rules which, together with the formal elements of the theory, would make the examples superfluous? To that question I have already suggested the following answer. The philosopher is at liberty to substitute rules for examples and, at least in principle, he can expect to succeed in doing so. In the process, however, he will alter the nature of the knowledge possessed by the community from which his examples were drawn. What he will be doing, in effect, is to substitute one means of data processing for another. Unless he is extraordinarily careful he will weaken the community's cognition by doing so. Even with care, he will change the nature of the community's future responses to some experimental stimuli.

Johnny’s education, though not in the science, provides a new sort of evidence for these claims. To identify swans, geese, and ducks by correspondence rules rather than by perceived similarity is to draw closed nonintersecting curves around each of the clusters in figure 2. What results is a simple Venn diagram, displaying three

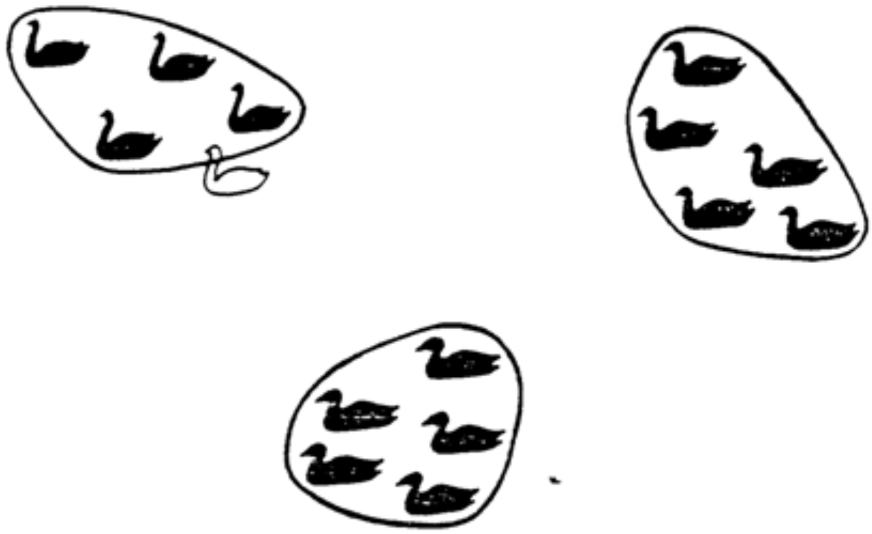

Figure 3

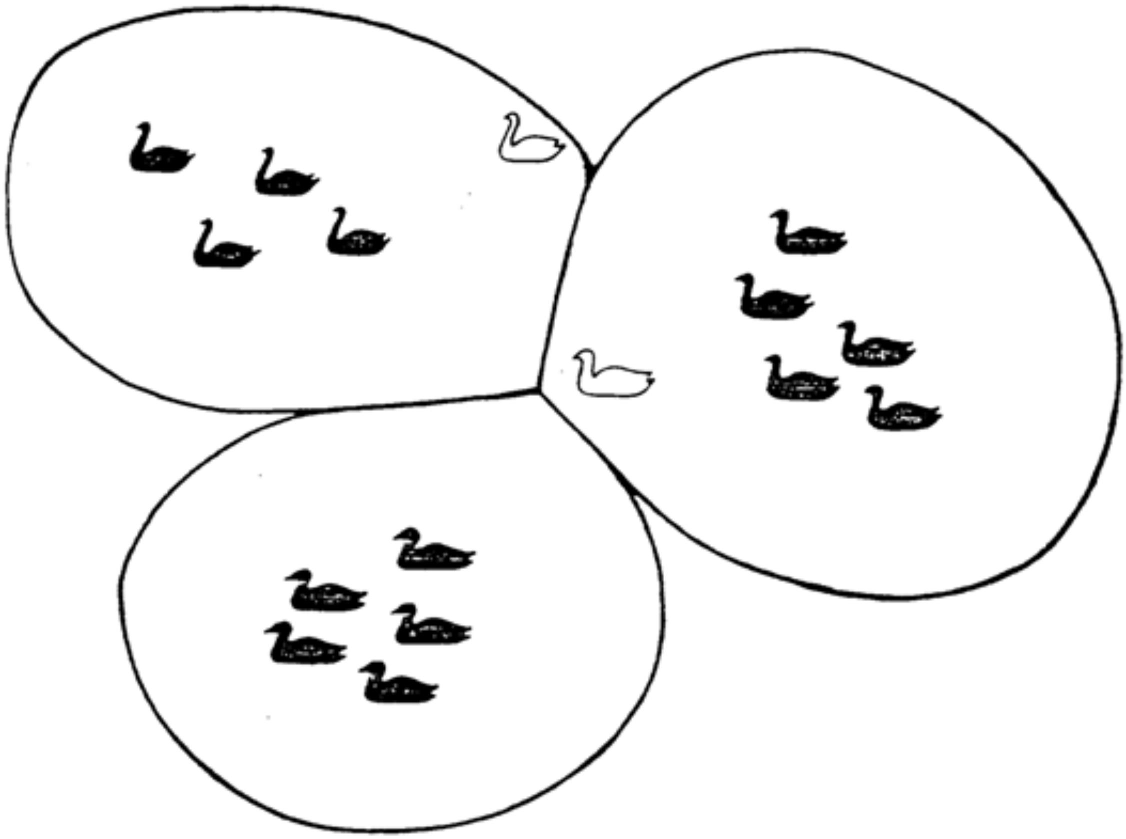

Figure 4

nonoverlapping classes. All swans lie in one, all geese in another, and so on. Where, however, should the curves be drawn? There are infinite possibilities. One of them is illustrated in figure 3, where boundaries are drawn very close to the bird figures in the three clusters. Given such boundaries, Johnny now can say what the criteria are for membership in the class of swans, geese, or ducks. On the other hand, he may be troubled by the very next waterfowl he sees. The outlined shape in the diagram is obviously a swan by the perceived distance criterion, but it is neither swan, goose, nor duck by the newly introduced correspondence rules for class membership.

Boundaries ought not, therefore, be drawn too near the edges of a cluster of exemplars. Let us therefore go to the other extreme, figure 4, and draw boundaries which exhaust most of the relevant parts of Johnny's perceptual space. With this choice, no bird that appears near one of the existing clusters will present a problem, but in avoiding that difficulty we have created another. Johnny used to know that there are no swan-geese. The new reconstruction of his knowledge deprives him of that information. Instead it supplies something he is extremely unlikely to need, the name that applies to a bird datum deep in the unoccupied space between swans and geese. To replace what has been lost we may imagine adding to John's cognitive apparatus a density function that describes the likelihood of his encountering a swan at various positions within the swan boundary, together with similar functions for geese and ducks. But the original similarity criterion supplied those already. In effect we would just have returned to the data-processing mechanism we had meant to replace.

Clearly, neither of the extreme techniques for drawing class boundaries will do. The compromise indicated in figure 5 is an obvious improvement. Any bird which appears near one of the existing clusters belongs to it. Any bird which appears midway between clusters has no name, but there is unlikely ever to be such datum. With class boundaries like these, Johnny should be able to operate successfully for some time. Yet he has gained nothing by substituting class boundaries for his original similarity criterion, and there has been some loss. If the strategic suitability of these boundaries is to be maintained, their location may need to be changed each time Johnny encounters another swan.

Figure 6 shows what I have in mind. Johnny has encountered one more swan. It lies, as it should, entirely within the old class boundary. There has been no problem of identification. But there may be one next time unless new boundaries, here shown as dotted lines, are drawn to take account of the altered shape of the swan cluster. Without the outward adjustment of the swan boundary, the very next bird encountered, though unproblematically a swan by the resemblance criterion, may fall on or even outside the old boundary. Without the simultaneous retraction of the duck boundary, the empty space, which Johnny's more experienced seniors have assured him can be preserved, would have become excessively narrow. If that is so if, that is, each new experience can demand some adjustment of the class boundaries, one may well ask whether Johnny was wise to allow philosophers to draw any such boundaries for him.

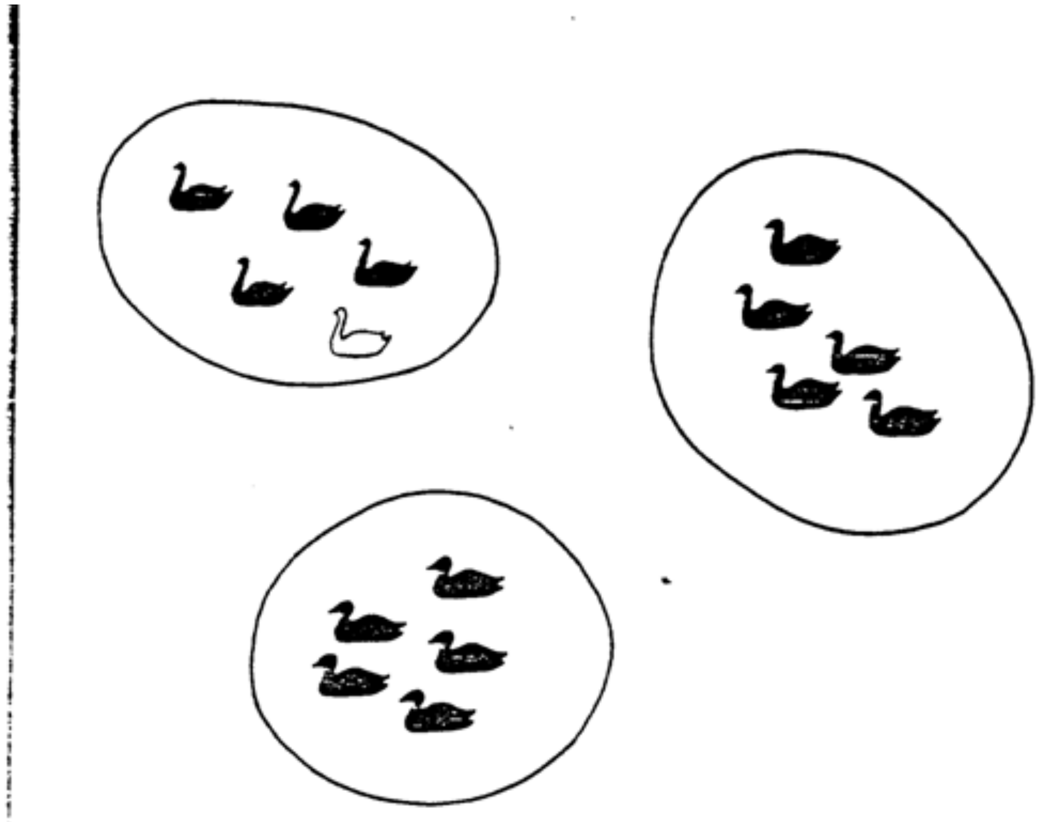

Figure 5

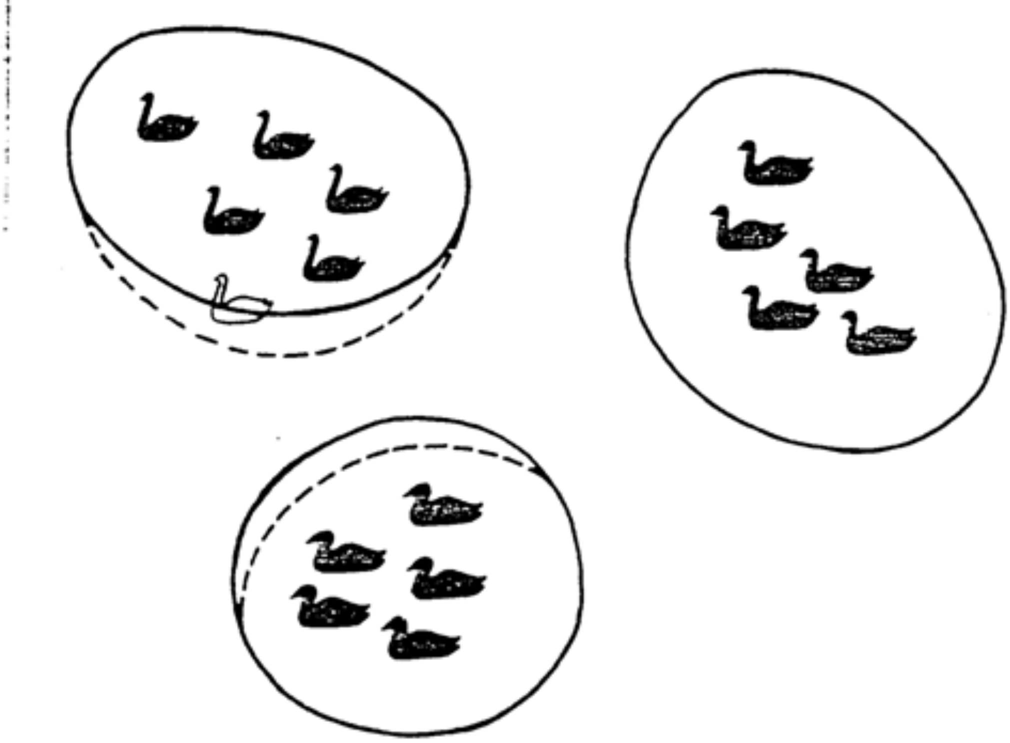

Figure 6

The primitive resemblance criterion he had previously acquired would have handled all these cases unproblematically and without continual adjustment. There is, I feel sure, such a thing as meaning change or change in the range of application of a term. But only the notion that meaning or applicability depends on predetermined boundaries could make us want to deploy any such phraseology here.

I am not, let me now emphasize, suggesting that there are never good reasons to draw boundaries or adopt correspondence rules. If Johnny had been presented with a series of birds that bridged the empty space between swans and geese, he would have been forced to resolve the resulting quandary with a line that divided the swan-goose continuum by definition. Or, if there were independent reasons for supposing that color is a stable criterion for the identification of waterfowl, Johnny might wisely have committed himself to the generalization, "all swans are white. " That strategy might save valuable data-processing time. In any case, the generalization would provide an entry point for logical manipulation. There are appropriate occasions for switching to the well-known strategy that relies upon boundaries and rules. But it is not the only available strategy for either stimuli- or data-processing. An alternative does exist, one based upon what I have been calling a learned perception of similarity. Observation, whether of language learning, scientific education, or scientific practice, suggests that it is, in fact, widely used. By ignoring it in epistemological discussion, we may do much violence to our understanding of the nature of knowledge.ered The Structure of Scientific Revolutions because I, the book's historian-author, could not, when examining the membership of a scientific community, retrieve enough shared rules to account for the group’s unproblematic conduct of research. Shared examples of successful practice could, I next concluded, provide what the group lacked in rules. Those examples were its paradigms,' and as such essential to its continued research. Unfortunately, having gotten that far, I allowed lowed the term's applications to expand, embracing all shared group commitments, all components of what I now wish to call the disciplinary matrix. Inevitably, the result was confusion, and it obscured the original reasons for introducing a special term. But those reasons still stand. Shared examples can serve cognitive functions commonly attributed to shared rules. When they do, knowledge develops differently from the way it does when governed by rules. This paper has, above all, been an effort to isolate, clarify, and drive home those essential points. If they can be seen, we shall be able to dispense with the term "paradigm," though not with the concept that led to its introduction.