hand, the second-order argument shows that R' has been successful each time and therefore demands that it not be trusted next time, i.e., calls for the prediction of 1. So the very definition of R' renders it impossible for the rule to be successful without being *incoherent*.[11] The suggested second-order argument in support of R' could be formulated only if R' were known to be unreliable, and would therefore be worthless. So we have an a priori reason for preferring R to its competitor R'. But it is easy to produce any number of alternative rules of inductive inference, none of which suffers from the fatal defect of R'. The choice between such rules, I suggest, has to be made in the light of experience of their use. I have tried to show in outline how such experience can properly be invoked without logical circularity.

[11] A parallel situation would arise in the use of R in predicting the members of the 1-0 series only if R were to be predominantly *unsuccessful*. But then we would have the best of reasons for assigning R zero strength, and the second-order argument would be pointless.

XIII

Models and Archetypes

SCIENTISTS often speak of using models but seldom pause to consider the presuppositions and the implications of their practice. I shall find it convenient to distinguish between a number of operations, ranging from the familiar and trivial to the farfetched but important, all of which are sometimes called "the use of models." I hope that even this rapid survey of a vast territory may permit a well-grounded verdict on the value of recourse to cognitive models.

억지의
무리한

To speak of "models" in connection with a scientific theory already smacks of the metaphorical. Were we called upon to provide a perfectly clear and uncontroversial example of a model, in the literal sense of that word, none of us, I imagine, would think of offering Bohr's model of the atom, or a Keynesian model of an economic system.

Typical examples of models in the literal sense of the word might include: the ship displayed in the showcase of a travel agency ("a model of the *Queen Mary*"), the airplane that emerges from a small boy's construction kit, the Stone Age village in the museum of natural history. That is to say, the standard cases are three-dimensional miniatures, more or less "true to scale," of some existing or imagined material object. It will be convenient to call the real or imaginary thing depicted by a model the *original* of that model.

① literal sense

miniature

model - original

We also use the word "model" to stand for a type of design (the dress designer's "spring models," the 1959 model Ford) —or to mean some exemplar (a model husband, a model solution of an equation). The senses in which a model is a type of design—or, on the other hand, something worthy of imitation—can usually be ignored in what follows.

② "model" to stand for a type of design

"1959 모델"

some examplar

이 둘은 무시하겠다

It seems arbitrary to restrict the idea of a model to something *smaller* than its original. A natural extension is to admit proportional change of scale in any relevant dimension

admit magnification, as in a larger-than-life-size likeness of a mosquito. A further natural extension is to admit proportional change of scale in *any* relevant dimension, such as time.

scale models

In all such cases, I shall speak of *scale models*. This label will cover all likenesses of material objects, systems, or processes, whether real or imaginary, that preserve relative proportions. They include experiments in which chemical or biological processes are artificially decelerated ("slow motion experiments") and those in which an attempt is made to imitate social processes in miniature.

slow motion experiments

The following points about scale models seem uncontroversial:

1. A scale model is always a model *of* something. The notion of a scale model is relational and, indeed, asymmetrically so: If *A* is a scale model of *B*, *B* is not a scale model of *A*.

2. A scale model is designed to serve a purpose, to be a means to some end. It is to show how the ship looks, or how the machine will work, or what law governs the interplay of parts in the original; the model is intended to be enjoyed for its own sake only in the limiting case where the hobbyist indulges a harmless fetishism.

3. A scale model is a representation of the real or imaginary thing for which it stands: its use is for "reading off" properties of the original from the directly presented properties of the model.

4. It follows that some features of the model are irrelevant or unimportant, while others are pertinent and essential, to the representation in question. There is no such thing as a perfectly faithful model; only by being unfaithful in *some* respect can a model represent its original.

5. As with all representations, there are underlying conventions of interpretation—correct ways for "reading" the model.

6. The conventions of interpretation rest upon partial identity of properties coupled with invariance of proportionality. In making a scale model, we try on the one hand

① feature

② relative proportion

to make it resemble the original by reproduction of some features (the color of the ship's hull, the shape and rigidity of the airfoil) and on the other hand to preserve the *relative* proportions between relevant magnitudes. In Peirce's terminology, the model is an *icon*, literally embodying the features of interest in the original.[1] It says, as it were: "*This* is how the original is."

icon

In making scale models, our purpose is to reproduce in a relatively manipulable or accessible embodiment, selected features of the "original": we want to see how the new house will look, or to find out how the airplane will fly, or to learn how the chromosome changes occur. We try to bring the remote and the unknown to our own level of middle-sized existence.[2]

reproduce
· manipulable
· accessible
features

There is, however, something self-defeating in this aim, since change of scale must introduce irrelevance and distortion. We are forced to replace living tissue by some inadequate substitute, and sheer change of size may upset the balance of factors in the original. Too small a model of a uranium bomb will fail to explode, too large a reproduction of a housefly will never get off the ground, and the solar system cannot be expected to look like its planetarium model. Inferences from scale model to original are intrinsically precarious and in need of supplementary validation and correction.

Icon:

[1] "An *Icon* is a sign which refers to the Object that it denotes merely by virtue of characters of its own, and which it possesses, just the same, whether any such Object actually exists or not. . . . Anything whatever . . . is an Icon of anything, in so far as it is like that thing and used as a sign of it." *Collected Papers of Charles Sanders Peirce* (Cambridge, Mass., 1931–35), II, 247.

[2] A good example of the experimental use of models is described in Victor P. Starr's article, "The General Circulation of the Atmosphere," *Scientific American*, CXCV (December 1956), 40–45. The atmosphere of one hemisphere is represented by water in a shallow rotating pan, dye being added to make the flow visible. When the perimeter of the pan is heated the resulting patterns confirm the predictions made by recent theories about the atmosphere.

2 analogue models

Let us now consider models involving *change of medium.* I am thinking of such examples as hydraulic models of economic systems, or the use of electrical circuits in computers. In such cases I propose to speak of *analogue models.*

reproduce the structure or web of relationship

An analogue model is some material object, system, or process designed to reproduce as faithfully as possible in some new medium the *structure* or web of relationships in an original. Many of our previous comments about scale models also apply to the new case. The analogue model, like the scale model, is a symbolic representation of some real or imaginary original, subject to rules of interpretation for making accurate inferences from the relevant features of the model.

The crucial difference between the two types of models is in the corresponding methods of interpretation. Scale models, as we have seen, rely markedly upon identity: their aim is to imitate the original, except where the need for manipulability enforces a departure from sheer reproduction. And when this happens the deviation is held to a minimum, as it were: geometrical magnitudes in the original are still *reproduced,* though with a constant change of ratio. On the other hand, the making of analogue models is guided by the more abstract aim of reproducing the *structure* of the original.

An adequate analogue model will manifest a point-by-point correspondence between the relations *it* embodies and those embodied in the original: every incidence of a relation in the original must be echoed by a corresponding incidence of a correlated relation in the analogue model. To put the matter in another way: there must be rules for translating the terminology applicable to the model in such a way as to conserve truth value.

Thus, the dominating principle of the analogue model is what mathematicians call "isomorphism."[3] We can, if we please, regard the analogue model as iconic of its original,

[3] For a more precise account of isomorphism, see for instance Rudolf Carnap, *Introduction to Symbolic Logic and Its Applications* (New York, 1958), p. 75.

as we did in the case of the scale model, but if we do so we must remember that the former is "iconic" in a more abstract way than the latter. The analogue model shares with its original not a set of features or an identical proportionality of magnitudes but, more abstractly, the same structure or pattern of relationships. Now identity of structure is compatible with the widest variety of content—hence the possibilities for construction of analogue models are endless.

analogue model

The remarkable fact that the same pattern of relationships, the same structure, can be embodied in an endless variety of different media makes a powerful and a dangerous thing of the analogue model. The risks of fallacious inference from inevitable irrelevancies and distortions in the model are now present in aggravated measure. Any would-be scientific use of an analogue model demands independent confirmation. Analogue models furnish plausible hypotheses, not proofs.

powerful but dangerous

I now make something of a digression to consider "mathematical models."[4] This expression has become very popular among social scientists, who will characteristically speak of "mapping" an "object system" upon one or another of a number of "mathematical systems or models."

3 mathematical models?

When used unemphatically, "model" in such contexts is often no more than a pretentious substitute for "theory" or "mathematical treatment." Usually, however, there are at least the following three additional suggestions: The original field is thought of as "projected" upon the abstract domain of sets, functions, and the like that is the subject matter of the correlated mathematical theory; thus social forces are said to be "modeled" by relations between mathematical entities. The "model" is conceived to be *simpler* and *more abstract* than the original. Often there is a suggestion of the model's being a kind of ethereal analogue model, as if the mathematical equations referred to an invisible mechanism whose

'theory' 'treatment' but. ① original field project (mathematical theory) abstract domain

[4] There is now a considerable literature on this subject. See Kenneth J. Arrow, "Mathematical Models in the Social Sciences," in D. Lerner, ed., *The Policy Sciences* (Stanford, Calif., 1951), pp. 129–154.

② model simpler more abstract

③ invisible mechanism

illusion

operation illustrates or even partially explains the operation of the original social system under investigation. This last suggestion must be rejected as an illusion.

The procedures involved in using a "mathematical model" seem to be the following:

1. In some original field of investigation, a number of relevant variables are identified, either on the basis of common sense or by reason of more sophisticated theoretical considerations. (For example, in the study of population growth we may decide that variation of population with time depends upon the number of individuals born in that time, the number dying, the number entering the area, and the number leaving.[5] I suppose these choices of variables are made at the level of common sense.)

2. Empirical hypotheses are framed concerning the imputed relations between the selected variables. (In population theory, common sense, supported by statistics, suggests that the numbers of births and deaths during any brief period of time are proportional both to that time and to the initial size of the population.)

3. Simplifications, often drastic, are introduced for the sake of facilitating mathematical formulation and manipulation of the variables. (Changes in a population are treated as if they were continuous; the simplest differential equations consonant with the original empirical data are adopted.)

③ simplification for formulation & manipulation

4. An effort is now made to solve the resulting mathematical equations—or, failing that, to study the *global* features of the mathematical systems constructed. (The mathematical equations of population theory yield the so-called "logistic function," whose properties can be specified completely. More commonly, the mathematical treatment of social data leads at best to "plausible topology," to use Kenneth Boulding's happy phrase;[6] i.e., qualitative conclusions con-

④ equation · system의 global feature

plausible topology

[5] Further details may be found conveniently in V. A. Kostitsyn, *Mathematical Biology* (London, 1939).

[6] "Economics as a Social Science," in *The Social Sciences at Mid-*

cerning distributions of maxima, minima, and so forth. This result is connected with the fact that the original data are in most cases at best *ordinal* in character.)

5. An effort is made to extrapolate to testable consequences in the original field. (Thus the prediction can be made that an isolated population tends toward a limiting size independent of the initial size of that population.)

testable

6. Removing some of the initial restrictions imposed upon the component functions in the interest of simplicity (e.g., their linearity) may lead to some increase in generality of the theory.

initial restriction

The advantages of the foregoing procedures are those usually arising from the introduction of mathematical analysis into any domain of empirical investigation, among them precision in formulating relations, ease of inference via mathematical calculation, and intuitive grasp of the structures revealed (e.g., the emergence of the "logistic function" as an organizing and mnemonic device).

mathematical analysis

The attendant dangers are equally obvious. The drastic simplifications demanded for success of the mathematical analysis entail a serious risk of confusing accuracy of the mathematics with strength of empirical verification in the original field. Especially important is it to remember that the mathematical treatment furnishes no *explanations*. Mathematics can be expected to do no more than draw consequences from the original empirical assumptions. If the functions and equations have a familiar form, there may be a background of pure mathematical research readily applicable to the illustration at hand. We may say, if we like, that the pure mathematics provides the *form* of an explanation, by showing what *kinds* of function would approximately fit the known data. But *causal* explanations must be sought elsewhere. In their inability to suggest explanations, "mathe-

trade off

③ no explanation

pure math → form of explanation

but no causal explanation

Century: Essays in Honor of Guy Stanton Ford (Minneapolis, 1952), p. 73.

matical models" differ markedly from the theoretical models now to be discussed.[7]

theoretical models.

In order now to form a clear conception of the scientific use of "theoretical models," I shall take as my paradigm Clerk Maxwell's celebrated representation of an electrical field in terms of the properties of an imaginary incompressible fluid. In this instance we can draw upon the articulate reflections of the scientist himself. Here is Maxwell's own account of his procedure:

> The first process therefore in the effectual study of the science must be one of simplification and reduction of the results of previous investigation to a form in which the mind can grasp them. The results of this simplification may take the form of a purely mathematical formula or of a physical hypothesis. In the first case we entirely lose sight of the phenomena to be explained; and though we may trace out the consequences of given laws, we can never obtain more extended views of the connexions of the subject. If, on the other hand, we adopt a physical hypothesis, we see the phenomena only through a medium, and are liable to that blindness to facts and rashness in assumption which a partial explanation encourages. We must therefore discover some method of investigation which allows the mind at every step to lay hold of a clear physical conception, without being committed to any theory founded on the physical science from which that conception is borrowed, so that it is neither drawn aside from the subject in pursuit of analytical subtleties, nor carried beyond the truth by a favourite hypothesis.[8]

Later comments of Maxwell's explain what he has in mind:

> By referring everything to the purely geometrical idea of the motion of an imaginary fluid, I hope to attain generality and precision, and to avoid the dangers arising from a premature

[7] It is perhaps worth noting that nowadays logicians use "model" to stand for an "interpretation" or "realization" of a formal axiom system. See John G. Kemeny, "Models of Logical Systems," *Journal of Symbolic Logic*, XIII (March 1948), 16–30.

[8] *The Scientific Papers of James Clerk Maxwell* (Cambridge University Press, 1890), I, 155–156.

> theory professing to explain the cause of the phenomena. . . . The substance here treated of . . . is not even a hypothetical fluid which is introduced to explain actual phenomena. It is merely a collection of imaginary properties which may be employed for establishing certain theorems in pure mathematics in a way more intelligible to many minds and more applicable to physical problems than that in which algebraic symbols alone are used.[9]

Points that deserve special notice are Maxwell's emphasis upon obtaining a "clear physical conception" that is both "intelligible" and "applicable to physical problems," his desire to abstain from "premature theory," and, above all, his insistence upon the "imaginary" character of the fluid invoked in his investigations. In his later elaboration of the procedure sketched above, the fluid seems at first to play the part merely of a mnemonic device for grasping mathematical relations more precisely expressed by algebraic equations held in reserve. The "exact mental image"[10] he professes to be seeking seems little more than a surrogate for facility with algebraic symbols.

Before long, however, Maxwell advances much farther toward ontological commitment. In his paper on action at a distance, he speaks of the "wonderful medium" filling all space and no longer regards Faraday's lines of force as "purely geometrical conceptions."[11] Now he says forthrightly that they "must not be regarded as mere mathematical abstractions. They are the directions in which the medium is exerting a tension like that of a rope, or rather, like that of our own muscles."[12] Certainly this is no way to talk about a collocation of imaginary properties. The purely geometrical medium has become very substantial.

A great contemporary of Maxwell is still more firmly committed to the realistic idiom. We find Lord Kelvin saying:

> We must not listen to any suggestion that we are to look upon the luminiferous ether as an ideal way of putting the thing. A

[9] *Ibid.*, I, 159–160. [10] *Ibid.*, II, 360. [11] *Ibid.*, II, 322. [12] *Ibid.*, II, 323.

real matter between us and the remotest stars I believe there is, and that light consists of real motions of that matter. . . . We know the luminiferous ether better than we know any other kind of matter in some particulars. We know it for its elasticity; we know it in respect to the constancy of the velocity of propagation of light for different periods. . . . Luminiferous ether must be a substance of most extreme simplicity. We might imagine it to be a material whose ultimate property is to be incompressible; to have a definite rigidity for vibrations in times less than a certain limit, and yet to have the absolutely yielding character that we recognize in wax-like bodies when the force is continued for a sufficient time.[13]

There is certainly a vast difference between treating the ether as a mere heuristic convenience, as Maxwell's first remarks require, and treating it in Kelvin's fashion as "real matter" having definite—though, to be sure, paradoxical—properties independent of our imagination. The difference is between thinking of the electrical field *as if* it were filled with a material medium, and thinking of it *as being* such a medium. One approach uses a detached comparison reminiscent of simile and argument from analogy; the other requires an identification typical of metaphor.

analogy vs metaphor (identification)

In *as if* thinking there is a willing suspension of ontological unbelief, and the price paid, as Maxwell insists, is absence of explanatory power. Here we might speak of the use of models as *heuristic fictions*. In risking existential statements, however, we reap the advantages of an explanation but are exposed to the dangers of self-deception by myths (as the subsequent history of the ether[14] sufficiently illustrates).

The *existential use of models* seems to me characteristic of the practice of the great theorists in physics. Whether we

[13] Sir William Thomson, *Baltimore Lectures* (London, 1904), pp. 8–12.

[14] See Sir Edmund Whittaker, *A History of the Theories of Aether and Electricity* (2nd ed.; London, 1951), I, especially chapter 9: "Models of the Aether." For further discussion of Maxwell's position, see Joseph Turner, "Maxwell on the Method of Physical Analogy," *British Journal for the Philosophy of Science*, VI (1955–56), 226–238.

consider Kelvin's "rude mechanical models,"[15] Rutherford's solar system, or Bohr's model of the atom, we can hardly avoid concluding that these physicists conceived themselves to be describing the atom *as it is*, and not merely offering mathematical formulas in fancy dress. In using theoretical models, they were not comparing two domains from a position neutral to both. They used language appropriate to the model in thinking about the domain of application: they worked not *by* analogy, but *through* and by means of an underlying analogy. Their models were conceived to be more than expository or heuristic devices.

Whether the fictitious or the existential interpretation be adopted, there is one crucial respect in which the sense of "model" here in question sharply diverges from those previously discussed in this paper. Scale models and analogue models must be actually put together: a merely "hypothetical" architect's model is nothing at all, and imaginary analogue models will never show us how things work in the large. But theoretical models (whether treated as real or fictitious) are not literally constructed: the heart of the method consists in *talking* in a certain way.

talking in a certain way.

It is therefore plausible to say, as some writers do, that the use of theoretical models consists in introducing a new language or dialect, suggested by a familiar theory but extended to a new domain of application. Yet this suggestion overlooks the point that the new idiom is always a description of some definite object or system (the model itself). If there is a change in manner of expression and representation, there is also the alleged depiction of a specific object or system, inviting further investigation.

The theoretical model need not be built; it is enough that it be *described*. But freedom to describe has its own liabilities. The inventor of a theoretical model is undistracted by accidental and irrelevant properties of the model object, which must have just the properties he assigns to it; but he is deprived of the controls enforced by the attempt at actual construction.

[15] Thomson, *op. cit.*, p. 12.

Even the elementary demand for self-consistency may be violated in subtle ways unless independent tests are available; and what is to be meant by the reality of the model becomes mysterious.

Although the theoretical model is described but not constructed, the sense of "model" concerned is continuous with the senses previously examined. This becomes clear as soon as we list the conditions for the use of a theoretical model.

1. We have an original field of investigation in which *some* facts and regularities have been established (in any form, ranging from disconnected items and crude generalizations to precise laws, possibly organized by a relatively well-articulated theory). ① original domain

2. A need is felt, either for explaining the given facts and regularities, or for understanding the basic terms applying to the original domain, or for extending the original corpus of knowledge and conjecture, or for connecting it with hitherto disparate bodies of knowledge—in short, a need is felt for further scientific mastery of the original domain. ② need

3. We describe some entities (objects, materials, mechanisms, systems, structures) belonging to a relatively unproblematic, more familiar, or better-organized secondary domain. The postulated properties of these entities are described in whatever detail seems likely to prove profitable. ③ familiar secondary domain

4. Explicit or implicit rules of correlation are available for translating statements about the secondary field into corresponding statements about the original field.

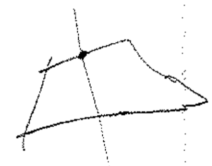
5. Inferences from the assumptions made in the secondary field are translated by means of the rules of correlation and then independently checked against known or predicted data in the primary domain. ⑤ secondary

The relations between the "described model" and the original domain are like those between an analogue model and its original. Here, as in the earlier case, the key to understanding the entire transaction is the identity of structure that in favorable cases permits assertions made about the

secondary domain to yield insight into the original field of interest.

Reliance upon theoretical models may well seem a devious and artificial procedure. Although the history of science has often shown that the right way to success is to "go round about" (as the Boyg advised Peer Gynt), one may well wonder whether the detour need be as great as it is in the use of models. Is the leap from the domain of primary interest to an altogether different domain really necessary? Must we really go to the trouble of using half-understood metaphors? Are the attendant risks of mystification and conceptual confusion unavoidable? And does not recourse to models smack too much of philosophical fable and literary allegory to be acceptable in a rational search for the truth? I shall try to show that such natural misgivings can be allayed.

The severest critic of the method will have to concede that recourse to models yields results. To become convinced of this, it is unnecessary to examine the great classical instances of large-scale work with models. The pragmatic utility of the method can be understood even more clearly in the simpler examples.



Consider, for instance, a recently published account of investigations in pure mathematics.[16] The problem to be solved was that of finding some method for dissecting any rectangle into a set of unequal squares—a problem of no practical importance, to be sure, and likely to interest only those who enjoy playing with "mathematical recreations." According to the authors' own account of their investigations, the direct path seemed to lead nowhere: trial and error (or "experiment," as they call it) and straightforward computation produced no results. The decisive breakthrough came when the investigators began to "go round about." As they put the matter: "In the next stage of the research we

[16] Martin Gardner (ed.), "Mathematical Games," *Scientific American*, CXCIX (November 1958), 136–142. The mathematicians were William T. Tutte, C. A. B. Smith, Arthur H. Stone, and R. L. Brooks.

$i = \frac{V}{r}$ $V = ir$

abandoned experiment for theory. We tried to represent rectangles by diagrams of different kinds. The last of these diagrams . . . suddenly made our problem part of the theory of electrical networks." [17]

Here we notice the deliberate introduction of a point-for-point model. Geometrical lines in the original figure were replaced by electrical terminals, squares by connecting wires through which electrical currents are imagined to flow. By suitable choices of the resistances in the wires and the strengths of the currents flowing through them, a circuit was described conforming to known electrical principles (Kirchoff's Laws). In this way, the resources of a well-mastered theory of electrical networks became applicable to the original geometrical problem. "The discovery of this electrical analogy," our authors say, "was important to us because it linked our problem with an established theory. We could now borrow from the theory of electrical networks and obtain formulas for the currents . . . and the sizes of the corresponding component squares." [18] This fascinating episode strikingly illustrates the usefulness of theoretical models.

It is sometimes said that the virtue of working with models is the replacement of abstractions and mathematical formulas by *pictures,* or any other form of representation that is readily visualized. But the example just mentioned shows that this view emphasizes the wrong thing. It is not easier to visualize a network of electrical currents than to visualize a rectangle dissected into component squares: the point of thinking about the electric currents is not that we can see or imagine them more easily, but rather that their properties are *better known* than those of their intended field of application. (And thus it makes perfectly good sense to treat something abstract, even a mathematical calculus, as a theoretical model of something relatively concrete.) To make good use of a model, we usually need intuitive grasp ("Gestalt knowledge") of its capacities, but so long as we can freely *draw inferences* from the model, its picturability is of no

[17] *Ibid.*, p. 136.

[18] *Ibid.*, p. 138.

importance. Whereas Maxwell turned away from the electrical field to represent it by a better-known model, subsequent progress in electrical theory now permits us to use the electrical field itself as a model for something else relatively unknown and problematical.

It has been said that the model must belong to a more "familiar" realm than the system to which it is applied. This is true enough, if familiarity is taken to mean belonging to a well-established and thoroughly explored theory. But the model need not belong to a realm of common experience. It may be as recondite as we please, provided we know how to use it. A promising model is one with implications rich enough to suggest novel hypotheses and speculations in the primary field of investigation. "Intuitive grasp" of the model means a ready control of such implications, a capacity to pass freely from one aspect of the model to another, and has very little to do with whether the model can literally be seen or imagined.

The case for the use of theoretical models is that the conditions favoring their success are sometimes satisfied; that sometimes it does prove feasible to invent models "better known" than the original subject matter they are intended to illuminate; and that it is often hard to conceive how the research in question could have been brought to fruition without recourse to the model. But there is also a formidable case against the use of theoretical models, which must now be heard.

Duhem

Nobody has attacked the use of models more eloquently or more savagely than the great French physicist Pierre Duhem. Here is a characteristic criticism:

> The French or German physicist conceives, in the space separating two conductors, abstract lines of force having no thickness or real existence, the English physicist materializes these lines and thickens them to the dimensions of a tube which he will fill with vulcanized rubber. In place of a family of ideal forces, conceivable only by reason, he will have a bundle of elastic strings, visible and tangible, firmly glued at both ends to the surfaces of

the two conductors, and, when stretched, trying both to contract and to expand. When the two conductors approach each other, he sees the elastic strings drawing closer together; then he sees each of them bunch up and grow large. Such is the famous model of electro-static action designed by Faraday and admired as a work of genius by Maxwell and the whole English school.[19]

Behind such passages as this is a conviction that the nineteenth-century English physicists were corrupting the ideals of science by abandoning clear definitions and a taut system of principles in logical array. "Theory is for him [the English physicist] neither an explanation nor a rational classification, but a model of these laws, a model not built for the satisfying of reason but for the pleasure of the imagination. Hence, it escapes the domination of logic." [20] Duhem might have tolerated, with a grimace, "those disparities, those incoherencies," [21] he disliked in the work of his English contemporaries could he have believed that models were fruitful. But he held them to be useless.

Oddly enough, Duhem applauds "the use of physical analogues" as "an infinitely valuable thing" and an altogether respectable "method of discovery." He is able to reconcile this approval with his strictures against models by purging reliance upon analogy of all its imaginative power. The two domains to be brought into relation by analogy must antecedently have been formulated as "abstract systems," and then, as he says, the demonstration of "an exact correspondence" will involve nothing "that can astonish the most rigorous logician." [22]

This is a myopic conception of scientific method; if much in scientific investigation offends the "rigorous logician,"

[19] *The Aim and Structure of Physical Theory*, trans. Philip P. Wiener (Princeton University Press, 1954), p. 70.

[20] *Ibid.*, p. 81.

[21] *Ibid.* Duhem took preference for working with models to be an expression of the English character. He thought the English, unlike the French, typically manifested "l'esprit de finesse" rather than "l'esprit géométrique."

[22] *Ibid.*, pp. 96–97.

model은 useless

physical analogues는 method of discovery로 정말 valuable?

abstract system

the truth may be that the rigor is out of place. To impose upon the exercise of scientific imagination the canons of a codified and well-ordered logical system is to run the risk of stifling research. Duhem's allegations of lack of coherence and clarity in the physical theories he was attacking must not be taken lightly. But this does not require us to treat the use of models as an aberration of minds too feeble to think about abstractions without visual aids.

It is instructive to compare Duhem's intemperate polemic with the more measured treatment of the same topic by a recent writer. In his valuable book, *Scientific Explanation*, Professor R. B. Braithwaite allows that "there are great advantages in thinking about a scientific theory through the medium of thinking about a model for it," but at once adds as his reason that "to do this avoids the complications and difficulties involved in having to think explicitly about the language or other form of symbolism by which the theory is represented." [23] That is to say that he regards the use of models as a *substitute* for the available alternative of taking the scientific theory "straight." The dominating notion in Braithwaite's conception of scientific theory is that of a "deductive scientific system" defined as "a set of hypotheses . . . arranged in such a way that from some of the hypotheses as premises all the other hypotheses logically follow." [24] The ideal form of scientific theory, for Braithwaite as for Duhem, is essentially that of Euclid's *Elements*—or, rather, Euclid as reformed by Hilbert. It is natural, accordingly, for Braithwaite to agree with Duhem in attaching little value to the use of models in science.

Braithwaite says that "the price of the employment of models is eternal vigilance"; [25] yet as much could be said for the employment of deductive systems or anything else. The crucial issue is whether the employment of models is to be regarded as a prop for feeble minds (as Duhem thought) or a convenient short cut to the consideration of deductive

[23] *Scientific Explanation* (Cambridge, 1953), p. 92.

[24] *Ibid.*, p. 12.

[25] *Ibid.*, p. 93.

(Braithwaite) convenient short cut로 model 유용

minor한

HD system

systems (as Braithwaite seems to think)—in short, as surrogate for some other procedure—or as a rational method having its own canons and principles. Should we think of the use of models as belonging to psychology—like doodles in a margin—or as having its proper place in the logic of scientific investigation? I have been arguing that models are sometimes not epiphenomena of research, but play a distinctive and irreplaceable part in scientific investigation: models are not disreputable understudies for mathematical formulas.

It may be useful to consider this central issue from another point of view. To many, the use of models in science has strongly resembled the use of metaphors. One writer says, "We are forced to employ models when, for one reason or another, we cannot give a direct and complete description in the language we normally use. Ordinarily, when words fail us, we have recourse to analogy and metaphor. The model functions as a more general kind of *metaphor*." [26]

Certainly there is some similarity between the use of a model and the use of metaphor—perhaps we should say, of a sustained and systematic metaphor. And the crucial question about the autonomy of the method of models is paralleled by an ancient dispute about the translatability of metaphors. Those who see a model as a mere crutch are like those who consider metaphor a mere decoration or ornament. But there are powerful and irreplaceable uses of metaphor not adequately described by the old formula of "saying one thing and meaning another." [27]

A memorable metaphor has the power to bring two separate domains into cognitive and emotional relation by using language directly appropriate to the one as a lens for seeing the other; the implications, suggestions, and supporting values entwined with the literal use of the metaphorical expression enable us to see a new subject matter in a new

[26] E. H. Hutten, "The Role of Models in Physics," *British Journal for the Philosophy of Science*, IV (1953–54), 289.

[27] For elaboration of this and related points, see Chapter III above.

way. The extended meanings that result, the relations between initially disparate realms created, can neither be antecedently predicted nor subsequently paraphrased in prose. We can comment *upon* the metaphor, but the metaphor itself neither needs nor invites explanation and paraphrase. Metaphorical thought is a distinctive mode of achieving insight, not to be construed as an ornamental substitute for plain thought.

Much the same can be said about the role of models in scientific research. If the model were invoked *after* the work of abstract formulation had already been accomplished, it would be at best a convenience of exposition. But the memorable models of science are "speculative instruments," to borrow I. A. Richards' happy title.[28] They, too, bring about a wedding of disparate subjects, by a distinctive operation of transfer of the *implications* of relatively well-organized cognitive fields. And as with other weddings, their outcomes are unpredictable. Use of a particular model may amount to nothing more than a strained and artificial description of a domain sufficiently known otherwise. But it may also help us to notice what otherwise would be overlooked, to shift the relative emphasis attached to details—in short, to *see new connections*.

speculative instrument.

to see new connection

A dissenting critic might be willing to agree that models are useful in the ways I have stated, and yet still harbor reservations about their rationality. "You have compared the use of models in science to the use of metaphors," I imagine him saying, "yet you cannot seriously contend that scientific investigation *requires* metaphorical language. That a model may lead to insight not otherwise attainable is just a fact of psychology. The *content* of the theory that finally emerges is wholly and adequately expressed by mathematical equations, supplemented by rules for co-ordination with the physical world. To count the model as an intrinsic part of the investigation is as plausible as including pencil sharpen-

[28] *Speculative Instruments* (London, 1955).

ing in scientific research. Your inflated claims threaten to debase the hard-won standards of scientific clarity and accuracy."

This objection treats the relation between the model and the formal theory by which it is eventually replaced as *causal;* it claims that the model is no more than a *de facto* contrivance for leading scientists to a deductive system. I cannot accept this view of the relation between model and theory. We have seen that the successful model must be isomorphic with its domain of application. So there is a rational basis for using the model. In stretching the language by which the model is described in such a way as to fit the new domain, we pin our hopes upon the existence of a common structure in both fields. If the hope is fulfilled, there will have been an objective ground for the analogical transfer. For we call a mode of investigation rational when it has a rationale, that is to say, when we can find reasons which justify what we do and that allow for articulate appraisal and criticism. The putative isomorphism between model and field of application provides such a rationale and yields such standards of critical judgment. We can determine the validity of a given model by checking the extent of its isomorphism with its intended application. In appraising models as good or bad, we need not rely on the sheerly pragmatic test of fruitfulness in discovery; we can, in principle at least, determine the "goodness" of their "fit."

We may deal with any residual qualms about the propriety of condoning metaphorical description in scientific research by stressing the limitations of any comparison between model and metaphor. The term "metaphor" is best restricted to relatively brief statements, and if we wished to draw upon the traditional terms of rhetoric we might better compare use of models with allegory or fable. But none of these comparisons will stand much strain.

Use of theoretical models resembles the use of metaphors in requiring analogical transfer of a vocabulary. Metaphor and model-making reveal new relationships; both are at-

tempts to pour new content into old bottles. But a metaphor operates largely with *commonplace* implications. You need only proverbial knowledge, as it were, to have your metaphor understood; but the maker of a scientific model must have prior control of a well-knit scientific theory if he is to do more than hang an attractive picture on an algebraic formula. Systematic complexity of the source of the model and capacity for analogical development are of the essence. As Stephen Toulmin says:

> It is in fact a great virtue of a good model that it does suggest further questions, taking us beyond the phenomena from which we began, and tempts us to formulate hypotheses which turn out to be experimentally fertile. . . . Certainly it is this suggestiveness, and systematic deployability, that makes a good model something more than a simple metaphor.[29]

I have tried to consider various senses of "model" in a systematic order, proceeding from the familiar construction of miniatures to the making of scale models in a more generalized way, and then to "analogue models" and "mathematical models," until we reached the impressive but mysterious uses of "theoretical models," where mere description of an imaginary but possible structure sufficed to facilitate scientific research. Now I propose to take one last step by considering cases where we have, as it were, an implicit or submerged model operating in a writer's thought. What I have in mind is close to what Stephen C. Pepper meant by "root metaphors." This is his explanation of the notion:

> The method in principle seems to be this: A man desiring to understand the world looks about for a clue to its comprehension. He pitches upon some area of common-sense fact and tries if he cannot understand other areas in terms of this one. The original area becomes then his basic analogy or root metaphor. He describes as best he can the characteristics of this area, or, if you will, discriminates its structure. A list of its structural characteristics becomes his basic concepts of explanation and description.

[29] *The Philosophy of Science* (London, 1953), pp. 38–39.

We call them a set of categories. In terms of these categories he proceeds to study all other areas of fact whether uncriticized or previously criticized. He undertakes to interpret all facts in terms of these categories. As a result of the impact of these other facts upon his categories, he may qualify and readjust the categories, so that a set of categories commonly changes and develops. Since the basic analogy or root metaphor normally (and probably at least in part necessarily) arises out of common sense, a great deal of development and refinement of a set of categories is required if they are to prove adequate for a hypothesis of unlimited scope. Some root metaphors prove more fertile than others, have greater power of expansion and adjustment. These survive in comparison with the others and generate the relatively adequate world theories.[30]

Pepper is talking about how metaphysical systems ("world hypotheses," as he calls them) arise; but his remarks have wider application. Use of a dominating system of concepts to describe a new realm of application by analogical extension seems typical of much theorizing:

> Any area for investigation, so long as it lacks prior concepts to give it structure and an express terminology with which it can be managed, appears to the inquiring mind inchoate—either a blank, or an elusive and tantalizing confusion. Our usual recourse is, more or less deliberately, to cast about for objects which offer parallels to dimly sensed aspects of the new situation, to use the better known to elucidate the less known, to discuss the intangible in terms of the tangible. This analogical procedure seems characteristic of much intellectual enterprise. There is a deal of wisdom in the popular locution for "what is its nature?" namely: "What's it *like?*" We tend to describe the nature of something in similes and metaphors, and the vehicles of these recurrent figures, when analyzed, often turn out to be the attributes of an implicit analogue through which we are viewing the object we describe.[31]

Here no *specific* structure or system is postulated by the theorist—there is not even a suppressed or implicit model.

[30] *World Hypotheses* (University of California Press, 1942), pp. 91–92.

[31] M. H. Abrams, *The Mirror and the Lamp* (Oxford University Press, 1953), pp. 31–32.

root metaphor | conceptual archetype (Kuhn's paradigm)

A system of concepts is used analogically, but there is no question of a definite explanation of given phenomena or laws. For reasons already given, I shall not follow Pepper in speaking of "metaphors." For want of a better term, I shall speak of "conceptual archetypes" or, more briefly, of "archetypes."[32] Others have perhaps had a similar idea in mind when they spoke of "ultimate frames of reference" or "ultimate presuppositions."

By an *archetype* I mean a systematic repertoire of ideas by means of which a given thinker describes, by *analogical extension,* some domain to which those ideas do not immediately and literally apply. Thus, a detailed account of a particular archetype would require a list of key words and expressions, with statements of their interconnections and their paradigmatic meanings in the field from which they were originally drawn. This might then be supplemented by analysis of the ways in which the original meanings become extended in their analogical uses.

A striking illustration of the influence of an archetype upon a theorist's work is to be found in the writings of Kurt Lewin. Ironically enough, he formally disclaims any intention of using models. "We have tried," he says, "to avoid developing elaborate 'models'; instead we have tried to represent the dynamic relations between the psychological facts by mathematical constructs at a sufficient level of generality."[33] Well, there may be no specific models envisaged; yet any reader of Lewin's papers must be impressed by the degree to which he employs a vocabulary indigenous to *physical* theory. We repeatedly encounter such words as "field," "vector," "phase-space," "tension," "force," "boundary," "fluidity"—visible symptoms of a massive archetype awaiting to be reconstructed by a sufficiently patient critic.

[32] The term is used in a rather different sense by literary critics as, for example, in Maud Bodkin's well-known *Archetypal Patterns in Poetry* (Oxford, 1934).

[33] Kurt Lewin, *Field Theory in Social Science* (New York, 1951), p. 21.

In this I see nothing to be deplored on the ground of general principles of sound method. Competent specialists must appraise the distinctive strengths and weaknesses of Lewin's theories; but an onlooker may venture to record his impression that Lewin's archetype, confused though it may be in detail, is sufficiently rich in implicative power to be a useful speculative instrument. It is surely no mere coincidence that Lewin's followers have been stimulated into making all manner of interesting empirical investigations that bear the stamp of their master's archetype. Now if an archetype is sufficiently fruitful, we may be confident that logicians and mathematicians will eventually reduce the harvest to order. There will always be competent technicians who, in Lewin's words, can be trusted to build the highways "over which the streamlined vehicles of a highly mechanized logic, fast and efficient, can reach every important point on fixed tracks." [34] But clearing intellectual jungles is also a respectable occupation. Perhaps every science must start with metaphor and end with algebra; and perhaps without the metaphor there would never have been any algebra.

Of course, there is an ever-present and serious risk that the archetype will be used metaphysically, so that its consequences will be permanently insulated from empirical disproof. The more persuasive the archetype, the greater the danger of its becoming a self-certifying myth. But a good archetype can yield to the demands of experience; while it channels its master's thought, it need not do so inflexibly. The imagination must not be confused with a strait jacket.

risk

If I have been on the right track in my diagnosis of the part played in scientific method by models and archetypes, some interesting consequences follow for the relations between the sciences and the humanities. All intellectual pursuits, however different their aims and methods, rely firmly upon such exercises of the imagination as I have been recalling. Similar archetypes may play their parts in different disciplines; a sociologist's pattern of thought may also be

[34] *Ibid.*, p. 3.

the key to understanding a novel. So perhaps those interested in excavating the presuppositions and latent archetypes of scientists may have something to learn from the industry of literary critics. When the understanding of scientific models and archetypes comes to be regarded as a reputable part of scientific culture, the gap between the sciences and the humanities will have been partly filled. For exercise of the imagination, with all its promise and its dangers, provides a common ground. If I have so much emphasized the importance of scientific models and archetypes, it is because of a conviction that the imaginative aspects of scientific thought have in the past been too much neglected. For science, like the humanities, like literature, is an affair of the imagination.