In Sidney Morgenbesser (Ed.), Philosophy of Science Today. New York: Basic Books, Inc., 1967. Pp. 55-67.

6 WHAT IS A SCIENTIFIC THEORY?

Patrick Suppes

Often when we ask what is a so-and-so, we expect a clear and definite answer. If, for example, someone asks me what is a rational number, I may give the simple and precise answer that a rational number is the ratio of two integers. There are other kinds of simple questions for which a precise answer can be given but for which in ordinary talk a rather vague answer is usually given and accepted. Someone reads about nectarines in a book, but has never seen a nectarine, or possibly has seen nectarines but is not familiar with their English name. He may ask me, "What is a nectarine?" and I would probably reply, "A smooth-skinned sort of peach." Certainly, this is not a very exact answer, but if my questioner knows what peaches are, it may come close to being satisfactory.

The kind of question I want to discuss fits neither one of these patterns. Scientific theories are not like rational numbers or nectarines. Certainly they are not like nectarines, for they are not simple physical objects. They are like rational numbers in not being physical objects, but they are totally unlike rational numbers in that scientific theories cannot be defined in any simple or direct way in terms of other non-physical, abstract objects.

Good examples of the kind of question we shall analyze in this chapter are provided by the familiar inquiries: "What is physics?," "What is psychology?," "What is science?" To none of these questions do we expect a simple and precise answer. On the other hand, there are many interesting things to be said about the sort of thing physics or psychology is. I shall be trying to show in this essay that this is also true of scientific theories.

THE STANDARD SKETCH

The standard sketch of scientific theories—and I emphasize the word "sketch"—runs something like the following. A scientific theory consists of two parts. One part is an abstract logical calculus. In addition to the vocabulary of logic, this calculus includes the primitive symbols of the theory, and the logical structure of the theory is fixed by stating the axioms or postulates of the theory in terms of its primitive symbols. For many theories the primitive symbols will be thought of as theoretical terms like "electron" or "particle" that are not possible to relate in any simple way to observable phenomena.

The second part of the theory is a set of rules that assign an empirical content to the logical calculus by providing what are usually called "co-ordinating definitions" or "empirical interpretations" for at least some of the primitive and defined symbols of the calculus. It is always emphasized that the first part alone is not sufficient to define a scientific theory, for without a systematic specification of the intended empirical interpretation of the theory, it is not possible in any sense to evaluate the theory as a part of science, although it can be studied simply as a piece of pure mathematics.

The most striking thing about this characterization is its highly schematic nature. Concerning the first part of a theory, the logical calculus, it is unheard of to find a substantive example of a theory actually worked out as a logical calculus in the writings of most philosophers of science. Much handwaving is indulged in to demonstrate that this working out of the logical calculus is simple in principle and only a matter of tedious detail, but concrete evidence is seldom given.

The sketch of the second part of a theory, that is, the co-ordinating definitions or empirical interpretations of some of the terms, is also highly schematic. A common defense of the relatively vague schema offered is that the variety of different empirical interpretations, for example, the many different methods of measuring mass—make a precise characterization difficult. Moreover, as we move from the precisely formulated theory on to the very loose and elliptical sort of experimental language used by almost all scientists, it is difficult to impose a definite pattern on the rules of empirical interpretation.

The view I want to support in this essay is not that this standard sketch is wrong, but rather that it is far too simple. Its very sketchiness makes it possible to omit both important properties of theories and significant distinctions that may be introduced between different theories.

MODELS VERSUS EMPIRICAL INTERPRETATIONS

To begin with, there has been a strong tendency on the part of many philosophers to speak of the first part of a theory as a logical calculus purely in syntactical terms. The co-ordinating definitions provided in the second part do not in the sense of modern logic provide an adequate semantics for the formal calculus. Quite apart from questions about direct empirical observations, it is pertinent and natural from a logical standpoint to talk about the models of the theory. These models are highly abstract, non-linguistic entities, often quite remote in their conception from empirical observations. It may well be asked what does the concept of a model have to add to the familiar discussions of empirical interpretation of theories.

I think it is true to say that most philosophers find it easier to talk about theories than about models of theories. The reasons for this are several, but perhaps the most important two are the following: In the first place, philosophers' examples of theories are usually quite simple in character, and therefore are easy to discuss in a straightforward linguistic manner. In the second place, the introduction of models of a theory inevitably introduces a stronger mathematical element into the discussion. It is a natural thing to talk about theories as linguistic entities—that is, to speak explicitly of the precisely defined set of sentences of the theory and the like—when the theories are given in what is called standard formalization.

Theories are ordinarily said to have a standard formalization when they are formulated within first-order logic. Roughly speaking, first-order logic is just the logic of sentential connectives and predicates holding for one type of object. Unfortunately, when a theory assumes more than first-order logic, it is neither natural nor simple to formalize it in this fashion. For example, if in axiomatizing geometry, we want to define lines as certain *sets* of points, we must work within a framework that already includes the ideas of set theory. To be sure, it is theoretically possible to axiomatize simultaneously geometry *and* the relevant portions of set theory, but this is awkward and unduly laborious.

Theories of more complicated structure like quantum mechanics, classical thermodynamics, or a modern quantitative version of learning theory, need to use not only general ideas of set theory but also many results concerning the real numbers. Formalization of such theories in first-order logic is utterly impractical. Theories of this sort are similar to the theories mainly studied in pure mathematics in their degree of complexity. In such contexts it is very much simpler to assert things about models of the theory rather than to talk directly and explicitly about the sentences of the theory, perhaps the main reason for this being that the notion of a sentence of the theory is not well defined when the theory is not given in standard formalization.

I would like to give just two examples in which the notion of model enters in a natural and explicit way in discussing scientific theories. The first example is concerned with the nature of measurement. The primary aim of a given theory of measurement is to show in a precise fashion how to pass from qualitative observations to the quantitative assertions needed for more elaborate theoretical stages of science. An analysis of how this passage from the qualitative to the quantitative may be performed is provided by axiomatizing appropriate algebras of experimentally realizable operations and relations. Given an axiomatized theory of measurement of some empirical quantity such as mass, distance, or force, the mathematical task is to prove a representation theorem for *models* of the theory which establishes, roughly speaking, that any empirical model is isomorphic to some numerical model of the theory. The existence of this isomorphism between models justifies the application of numbers to things.

We cannot literally take a number in our hands and apply it to a physical object. What we can do is to show that the structure of a set of phenomena under certain empirical operations is the same as the structure of some set of numbers under arithmetical operations and relations. The definition of isomorphism of models in the given context makes the intuitive idea of *same structure* precise. The great significance of finding such an isomorphism of models is that we may then use all our familiar knowledge of computational methods, as applied to the arithmetical model, to infer facts about the isomorphic empirical model. A linguistic formulation of this central notion of an empirical model of a theory of measurement being isomorphic to a numerical model is extremely awkward and tedious to formulate. But in model-theoretic terms the notion is simple and in fact represents a direct application of the very general notion of isomorphism used throughout all domains of pure mathematics.

The second example of the use of models concerns the discussion of reductionism in the philosophy of science. Many of the problems formulated in connection with the question of reducing one science to another may be formulated as a series of problems using the notion of a representation theorem for the models of a theory. For instance, the thesis that psychology may be reduced to physiology would be for many people appropriately established if one could show that for any model of a psychological theory it was possible to construct an isomorphic model within physiological theory.

The absence at the present time of any deep unitary theory within either psychology or physiology makes present attempts to settle such a question of reductionism rather hopeless. The classical example from physics is the reduction of thermodynamics to statistical mechanics. Although this reduction is usually not stated in absolutely satisfactory form from a logical standpoint, there is no doubt that it is substantially correct and represents one of the great triumphs of mathematical physics.

INTRINSIC VERSUS EXTRINSIC CHARACTERIZATION

Quite apart from the two applications just mentioned of the concept of a model of a theory, we may bring this concept to bear directly on the question of characterizing a scientific theory. The contrast I wish to draw is between intrinsic and extrinsic characterization. The formulation of a theory as a logical calculus or, to put it in terms that I prefer, as a theory with a standard formalization, gives an intrinsic characterization, but this is certainly not the only approach. For instance, a natural question to ask within the context of logic is if a certain theory *can* be axiomatized with standard formalization, that is, within first-order logic. In order to formulate such a question in a precise manner, it is necessary to have some extrinsic way of characterizing the theory. One of the simplest ways of providing such an extrinsic characterization is simply to define the intended class of models of the theory. To ask if we can axiomatize the theory is then just to ask if we can state a set of axioms such that the models of these axioms are precisely the models in the defined class.

As a very simple example of a theory formulated both extrinsically and intrinsically, consider the extrinsic formulation of the theory of simple orderings that are isomorphic to a set of real numbers under the familiar less-than relation. That is, consider the class of all binary relations isomorphic to some fragment of the less-than relation for the real numbers. The extrinsic characterization of a theory usually follows the sort given for these orderings: namely, we designate a particular model of the theory (in this case, the numerical less-than relation) and then characterize the entire class of models of the theory in relation to this distinguished model.

The problem of intrinsic characterization is now to formulate a set of axioms that will characterize this class of models without referring to the relation between models, but only to the intrinsic properties of any one model. With the present case the solution is relatively simple, although even it is not easily formulated within first-order logic. The intrinsic axioms are just those for a simple ordering plus the axiom that the ordering must contain in its domain a countable subset, dense with respect to the ordering in question.

A casual inspection of scientific theories suggests that the usual formulations are intrinsic rather than extrinsic in character, and therefore that the question of extrinsic formulations usually arises only in pure mathematics. This would also seem to be a happy result, for our philosophical intuition is surely that an intrinsic characterization is in general to be preferred to an extrinsic one.

However, the problem of intrinsic axiomatization of a scientific theory is more complicated and considerably more subtle than this remark would indicate. Fortunately, it is precisely by explicit consideration of the class of models of the theory that the problem can be put into proper perspective and formulated in a fashion that makes possible consideration of its exact solution. I shall give one simple example. The axioms for classical particle mechanics are ordinarily stated in such a way that a co-ordinate system, as a frame of reference, is tacitly assumed.

One effect of this is that relationships deducible from the axioms are not necessarily invariant with respect to Galilean transformations. We can view the tacit assumption of a frame of reference as an extrinsic aspect of the familiar characterizations of the theory. From the standpoint of the models of the theory, the difficulty in the standard axiomatizations of mechanics is that a large number of formally distinct models may be used to express the same mechanical facts. Each of these different models represents the tacit choice of a different frame of reference, but all models representing the same mechanical facts are related by Galilean transformations. It is thus fair to say that in this instance the difference between models related by Galilean transformations does not have any theoretical significance, and it may be regarded as a defect of the axioms that these trivially distinct models exist. It is important to realize that this point about models related by Galilean transformations is not the kind of point usually made under the heading of empirical interpretations of the theory.

It is a conceptual point that just as properly belongs to the theoretical side of physics. I have introduced this example here in order to provide a simple instance of how the explicit consideration of models can lead to a more subtle discussion of the nature of a scientific theory. It is certainly possible from a philosophical standpoint to maintain that particle mechanics as a scientific theory should be expressed only in terms of Galilean invariant relationships, and that the customary formulations are defective in this respect.

CO-ORDINATING DEFINITIONS AND HIERARCHY OF THEORIES

I turn now to the second part of theories mentioned above. It is true that in the foregoing discussion we have been using the word "theory" to refer only to the first part of theories—that is, to the axiomatization of the theory or the expression of the theory as a logical calculus—but as I emphasized at the beginning, the necessity of providing empirical interpretation of a theory is just as important as the development of the formal side of the theory. My central point on this aspect of theories is that the story is much more complicated than the familiar remarks about co-ordinating definitions and empirical interpretations of theories would indicate. The kind of co-ordinating definitions often described by philosophers have their place in popular philosophical expositions of theories, but in the actual practice of testing scientific theories a more elaborate and more sophisticated formal machinery for relating a theory to data is required.

The concrete experience that scientists label an experiment cannot itself be connected to a theory in any complete sense. That experience must be put through a conceptual grinder that in many cases is excessively coarse. Once the experience is passed through the grinder, often in the form of the quite fragmentary records of the complete experiment, the experimental data emerge in canonical form and constitute a model of the experiment. It is this model of the experiment rather than a model of the theory for which direct co-ordinating definitions are provided. It is also characteristic that the model of the experiment is of a relatively different logical type from that of the model of the theory. It is common for the models of a theory to contain continuous functions or infinite sequences, but for the model of the experiment to be highly discrete and finitistic in character.

The assessment of the relation between the model of the experiment and some designated model of the theory is a characteristic fundamental problem of modern statistical methodology. What is important about this methodology for present purposes is that, in the first place, it is itself formal and theoretical in nature; secondly, it has been a typical function of this methodology to develop an elaborate theory of experimentation that intercedes between any fundamental scientific theory and raw experimental experience.

It is not possible in the confines of this essay to spell out the rather elaborate hierarchy of theories that are customarily interposed between the fundamental scientific theory and the experiments presumed to support it. My only point here is to make explicit the existence of this hierarchy and to point out that there is no simple procedure for giving co-ordinating definitions for a theory. It is even a bowdlerization of the facts to say that co-ordinating definitions are given to establish the proper connections between models of the theory and models of the experiment in the sense of the canonical form of the data just mentioned. The elaborate methods, for example, for estimating theoretical parameters in the model of the theory from models of the experiment are not adequately covered by a reference to co-ordinating definitions.

If someone asks, "What is a scientific theory?" it seems to me there is no simple response to be given. Are we to include as part of the theory the well-worked-out statistical methodology for testing the theory? If we are to take seriously the standard claims that the co-ordinating definitions are part of the theory, then it would seem inevitable that we must also include in a more detailed description of theories a methodology for designing experiments, estimating parameters and testing goodness-of-fit of the models of the theory. It does not seem to me important to give precise definitions of the form: X is a scientific theory if, and only if, so-and-so. What is important is to recognize that the existence of a hierarchy of theories arising from the methodology of experimentation for testing the fundamental theory is an essential ingredient of any sophisticated scientific discipline.

INSTRUMENTAL VIEW OF THEORIES

There is one view of scientific theories which is undoubtedly of considerable importance and which I have not yet mentioned. This is the view that theories are to be looked at from an instrumental viewpoint. The most important function of a theory, according to this view, is not to organize or assert statements that are true or false but to furnish material principles of inference that may be used in inferring one set of facts from another. Thus, in the familiar syllogism "all men are mortal; Socrates is a man; therefore, Socrates is mortal," the major premise "all men are mortal," according to this instrumental viewpoint, is converted into a principle of inference. And the syllogism now has only the minor premise "Socrates is a man."

From a logical standpoint it is clear that this is a fairly trivial move, and the question naturally arises if there is anything more substantial to be said about the instrumental viewpoint. Probably the most interesting argument for claiming that there is more than a verbal difference between these two ways of looking at theories or laws is the argument that when theories are regarded as principles of inference rather than as major premises, we are no longer concerned directly to establish their truth or falsity but to evaluate their *usefulness* in inferring new statements of fact. It is characteristic of discussions in this vein by philosophers that no genuinely original formal notions have arisen out of these discussions to displace the classical semantical notions of truth and validity. To talk, for instance, about laws having different jobs than statements of fact is trivial unless some systematic semantical notions are introduced to replace the standard analysis.

From another direction there has been one concerted serious effort to provide a formal framework for the evaluation of theories which replaces the classical concept of truth. What I have in mind is modern statistical decision theory. It is typical of statistical decision theory to talk about actions rather than statements. Once the focus is shifted from statements to actions, it seems quite natural to replace the concept of truth by that of expected loss or risk. It is appropriate to ask if a statement is true, but it does not make much sense to ask if it is risky. On the other hand, it is reasonable to ask how risky an action is, but not to ask if it is true. It is apparent that statistical decision theory, when taken literally, projects a more radical instrumental view of theories than does the view already sketched.

Theories are not regarded even as principles of inference but as methods of organizing evidence to decide which one of several actions to take. When theories are regarded as principles of inference, it is a straightforward matter to return to the classical view and to connect a theory as a principle of inference with the concept of a theory as a true major premise in an argument. The connection between the classical view and the view of theories as instruments leading to the taking of an action is certainly more remote and indirect.

Although many examples of applications of the ideas of statistical decision theory have been worked out in recent literature on the foundations of statistics, these examples in no case deal with complicated scientific theories, and I have seen no serious discussion of the treatment of scientific theories from the standpoint of statistical decision theory. Again, it is fair to say that when we want to talk about the evaluation of a sophisticated scientific theory, disciplines like statistical decision theory have not yet offered any genuine alternative to the semantical notions of truth and validity. In fact, even a casual inspection of the literature of statistical decision theory shows that in spite of the instrumental orientation of the fundamental ideas, formal development of the theory is wholly dependent on the standard semantical notions and in no sense replaces them.

What I mean by this is that in concentrating on the taking of an action as the terminal state of an inquiry the decision theorists have found it necessary to use standard semantical notions in describing evidence, their own theory, and so forth. For instance, I cannot recall a single discussion by decision theorists in which particular observation statements are treated in terms of utility rather than in terms of their truth or falsity.

It seems apparent that statistical decision theory does not at the present time offer a genuinely coherent or deeply original new view of scientific theories. Perhaps future developments of decision theory will proceed in this direction. Be that as it may, there is one still more radical instrumental view that I would like to discuss as the final point to be covered in this essay. As I have already noted, it is characteristic of many instrumental analyses to distinguish the status of theories from the status of particular assertions of fact. It is the point of a more radical instrumental, behavioristic view of the use of language to challenge this distinction and to look at the entire use of language, including the statement of theories as well as of particular matters of fact, from a behavioristic viewpoint.

According to this view of the matter, all uses of language are to be analyzed with strong emphasis on the language users. It is claimed that the semantical analysis of modern logic gives a very inadequate account even of the cognitive uses of language, because it does not explicitly consider the production and reception of linguistic stimuli by speakers, writers, listeners, and readers. It is plain that for the behaviorist an ultimately meaningful answer to the question "What is a scientific theory?" cannot be given in terms of the kinds of concepts considered earlier. An adequate and complete answer can be given only in terms of an explicit and detailed consideration of both the producers and consumers of the theory. There is much that is attractive in this behaviorist way of looking at theories or language in general. What it lacks at present, however, is sufficient scientific depth and definiteness to serve as a genuine alternative to the precise notions of modern logic and semantics. Moreover, much of the language of models and theories discussed earlier in this chapter is surely so approximately correct that any behaviorist revision of our way of looking at theories must yield the ordinary talk about models and theories as a first approximation. It is a matter for the future to see whether or not the behaviorist's approach will deepen our understanding of the nature of scientific theories.

In current perspective, the methods and concepts of modern logic provide a satisfactory and powerful set of tools for analyzing the detailed structure of scientific theories. What would seem to be needed for the present is deeper and more detailed application of these tools to the job of analysis. I have tried to indicate what I think are some of the more fruitful directions for future investigation.