On the "Standard Conception" of Scientific Theories

1. Theories: A Preliminary Characterization

Theories, it is generally agreed, are the keys to the scientific understanding of empirical phenomena: to claim that a given kind of phenomenon is scientifically understood is tantamount to saying that science can offer a satisfactory theoretical account of it.¹

Theories are normally constructed only when prior research in a given field has yielded a body of knowledge that includes empirical generalizations or putative laws concerning the phenomena under study. A theory then aims at providing a deeper understanding by construing those phenomena as manifestations of certain underlying processes governed by laws which account for the uniformities previously studied, and which, as a rule, yield corrections and refinements of the putative laws by means of which those uniformities had been previously characterized.

Prima facie, therefore, the formulation of a theory may be thought of as calling for statements of two kinds; let us call them internal principles and bridge principles for short. The internal principles serve to characterize the theoretical setting or the "theoretical scenario": they specify the basic entities and processes posited by the theory, as well as the theoretical laws that are assumed to govern them. The bridge principles, on the other hand, indicate the ways in which the scenario is linked to the previously examined phenomena which the theory is intended to explain. This general conception applies equally, I think, to the two types of theory which Nagel, following Rankine, distinguishes in his thorough study of the subject.² namely, "abstractive" theories, such as the Newtonian theory of gravitation and motion, and "hypothetical" theories, such as the kinetic theory of heat or the undulatory and corpuscular theories of light.

If I and B are the sets of internal and bridge principles by which a theory T is characterized, then T may be represented as the ordered couple of those sets:

(1a) \[ T = (I, B) \]

Or alternatively, and with greater intuitive appeal, T may be construed as the set of logical consequences of the sum of the two sets:

(1b) \[ T = c(IUB) \]

The formulation of the internal principles will typically make use of a theoretical vocabulary \( V_T \), i.e., a set of terms not employed in the earlier descriptions of, and generalizations about, the empirical phenomena which T is to explain, but rather introduced specifically to characterize the theoretical scenario and its laws. The bridge principles will evidently contain both the terms of \( V_T \) and those of the vocabulary used in formulating the original descriptions of, and generalizations about, the phenomena for which the theory is to account. This vocabulary will thus be available and understood before the introduction of the theory, and its use will be governed by principles which, at least initially, are independent of the theory. Let us refer to it as the pre-theoretical, or antecedent, vocabulary, \( V_A \), relative to the theory in question.

The antecedently examined phenomena for which a theory is to account have often been conceived as being described, or at least describable, by means of an observational vocabulary, i.e., a set of terms standing for particular individuals or for general attributes which, under suitable conditions, are accessible to "direct observation" by human observers. But this conception has been found inadequate on several important counts.³

The distinction I have suggested between theoretical and antecedent


vocabularies hinges on no such assumption. The terms of the antecedent vocabulary need not, and indeed should not, generally be conceived as observational in the narrow sense just adumbrated, for the antecedent vocabulary of a given theory will often contain terms which were originally introduced in the context of an earlier theory, and which are not observational in a narrow intuitive sense. Let us look at some examples.

In the classical kinetic theory of gases, the internal principles are assumptions about the gas molecules; they concern their size, their mass, their large number; and they include also various laws, partly taken over from classical mechanics, partly statistical in nature, pertaining to the motions and collisions of the molecules, and to the resulting changes in their momenta and energies. The bridge principles include statements such as that the temperature of a gas is proportional to the mean kinetic energy of its molecules, and that the rates at which different gases diffuse through the walls of a container are proportional to the numbers of molecules of the gases in question and to their average speeds. By means of such bridge principles, certain micro-characteristics of a gas, which belong to the scenario of the kinetic theory, are linked to macroscopic features such as temperature, pressure, and diffusion rate; these can be described, and generalizations concerning them can be formulated, in terms of an antecedently available vocabulary, namely, that of classical thermodynamics. And some of the features in question might well be regarded as rather directly observable or measurable.

Take, on the other hand, the theoretical account that Bohr’s early theory of the hydrogen atom provided for certain previously established empirical laws, such as these: the light emitted by glowing hydrogen gas is limited to certain characteristic discrete wavelengths, which correspond to a set of distinct lines in the emission spectrum of hydrogen; these wavelengths conform to certain general mathematical formulas, the first and most famous of which was Balmer’s

$$\lambda = b \frac{n^2}{n^2 - 4}$$

Here, $b$ is a numerical constant; and when $n$ is given the values $3, 4, 5, \ldots$, the formula yields the wavelengths of the lines that form the so-called Balmer series in the spectrum of hydrogen.

"STANDARD CONCEPTION" OF SCIENTIFIC THEORIES

Now let us look briefly at the internal principles and the bridge principles of the theory by which Bohr explained these and other empirical laws concerning the hydrogen spectrum.

The internal principles formulate Bohr’s conception that a hydrogen atom consists of a nucleus about which an electron circles in one or another of a set of discrete orbits with radii $r_1, r_2, r_5, \ldots$, where $r_1$ is proportional to $r_2^2$; that when the electron is in the $i$th orbit, it has an energy $E_i$ characteristic of that orbit and proportional to $(-1/r_i)$; that the electron can jump from a narrower to a wider orbit, or vice versa, and that in this process it absorbs or emits an amount of energy that equals the absolute difference between the energies associated with those orbits.

The bridge principles, which connect these goings-on with the optical phenomena to be explained, include statements such as these: (a) the light given off by glowing hydrogen gas results from the emission of energy by those atoms whose electrons happen to be jumping from outer to inner orbits; (b) the energy released by an electron jump from the $i$th to the $j$th orbit ($i > j$) is given off in the form of monochromatic electromagnetic waves with the wavelength $\lambda = (h \cdot c)/(E_i - E_j)$, where $h$ is Planck’s constant and $c$ the velocity of light.

As is to be expected, these bridge principles contain, on the one hand, certain theoretical terms such as ‘electronic orbit’ and ‘electron jump,’ which were specifically introduced to describe the theoretical scenario; on the other hand, they contain also certain antecedently available terms, such as ‘hydrogen gas,’ ‘spectrum,’ ‘wavelength of light,’ ‘velocity of light,’ and ‘energy.’ And clearly at least some of these terms—for example, ‘wavelength of light’ and ‘hydrogen gas’—are not observational terms in the intuitive sense mentioned earlier. Nonetheless, the terms are antecedently understood in the sense indicated above; for when Bohr proposed his theory of the hydrogen atom, principles for their use, including principles for the measurement of optical wavelengths, were already available; they were based on antecedent theories, including wave optics.

2. The Construal of Theories as Interpreted Calculi

In the analytic philosophy of science, theories have usually been char-
Carl G. Hempel

caracterized in a manner rather different from the one just outlined; and, at least until recently, this characterization was so widely accepted that it could count as the “standard,” or the “received,” philosophical construal of scientific theories.⁴ On this construal, too, a theory is characterized by two constituents, which, moreover, have certain clear affinities to what were called above its internal principles and its bridge principles.

The first constituent is an axiomatized deductive system—sometimes referred to as a calculus—of uninterpreted formulas, the postulates of the system corresponding to the basic principles of the theory. Thus, roughly speaking, the postulates of the calculus may be thought of as formulas obtained by axiomatizing the internal principles of the theory and then replacing the primitive theoretical terms in the axioms by variables or by dummy constants.

The second constituent is a set of sentences that give empirical import or applicability to the calculus by interpreting some of its formulas in empirical terms—namely, in terms of the vocabulary that serves to describe the phenomena which the theory is to explain. These sentences, which evidently are akin to the bridge principles mentioned above, were characterized by Campbell and by Ramsey as forming a “dictionary” that relates the theoretical terms to pre-theoretical ones;⁵ other writers have referred to them as “operational definitions” or “coordinative definitions” for the theoretical terms, as “rules of correspondence,” or as “interpretative principles.”

The standard conception, then, may be schematized by representing a theory as an ordered couple of sets of sentences:

\[ T = (C, R) \]


³ “STANDARD CONCEPTION” OF SCIENTIFIC THEORIES

where C is the set of formulas of the calculus and R the set of correspondence rules.

Whereas the bridge principles invoked in our initial characterization of a theory are conceived as a subset of the class of sentences asserted by the theory, the status of the correspondence rules in the standard construal is less clear. One plausible construal of them would be as terminological rules belonging to the metalanguage of the theory, which stipulate the truth, by definition or more general terminological convention, of certain sentences (in the language of the theory) that contain both theoretical and pre-theoretical terms. For this reason, no immediate analogue to (1b) is available as an alternative schematization of the standard view. The status of the correspondence rules will be examined further in section 6 below.

It would be a task of interest both for the history and for the philosophy of science to locate the origins of the standard conception and to trace its development in some detail. Such a study would surely have to take account of Reichenbach’s characterization of physical geometry (i.e., the theory of the geometrical structure of physical space) as an abstract, uninterpreted system of “pure” or mathematical geometry, supplemented by a set of coordinative definitions for the primitives,⁶ and it would have to consider Poincaré’s and Einstein’s views on the geometrical structure of physical space.

Campbell and some other proponents of the standard conception make provision for a third constituent of a theory—Campbell calls it an analogy, others (Nagel among them) call it a model—which is said to characterize the basic ideas of the theory by means of concepts with which we are antecedently acquainted, and which are governed by familiar empirical laws that have the same form as some of the basic principles of the theory. The role of models in this sense will be considered later; until then, the standard conception of theories will be understood in the sense of schema (2). I have myself relied on the standard construal in several earlier studies,⁷ but I have now come to consider it misleading in certain philosophically significant respects, which I will try to indicate in the following sections.

⁴ This idea is set forth very explicitly in chapter 8 of H. Reichenbach, The Rise of Scientific Philosophy (Berkeley and Los Angeles: University of California Press, 1951).

Carl G. Hempel

3. The Role of an Axiomatized Calculus in the Formulation of a Theory

My misgivings do not concern the obvious fact that theories as actually stated and used by scientists are almost never formulated in accordance with the standard schema; nor do they stem from the thought that a standard formulation could at best represent a theory quick-frozen, as it were, at one momentary stage of what is in fact a continually developing system of ideas. These observations represent no telling criticisms, I think, for the standard construal was never claimed to provide a descriptive account of the actual formulation and use of theories by scientists in the ongoing process of scientific inquiry; it was intended, rather, as a schematic explication that would clearly exhibit certain logical and epistemological characteristics of scientific theories.

This defense of the standard conception, however, naturally suggests this question: What are the logical and epistemological characteristics of theories that schema (2) serves to exhibit and illuminate? Let us consider in turn the various features which the schema attributes to a theory, beginning with the axiomatized calculus.

What is to be said in support of assuming axiomatization? It might quite plausibly be argued that an axiomatic exposition is an indispensable device for an unambiguous statement of a theory. For a theory has to be conceived as asserting a set of sentences that is closed under the relation of logical consequence in the sense that it contains all logical consequences (expressible in the language of the theory) of any of its subsets. A theory will therefore amount to an infinite set of sentences. In order to specify unambiguously the infinite set of sentences that a proposed theory is intended to assert, it will be necessary to provide a general criterion determining, for any sentence S, whether S is asserted by the theory. Axiomatization yields such a criterion: S is asserted by the theory just in case S is deducible from the specified axioms, or postulates.

This criterion determines membership in the intended set of sentences unambiguously, but it does not provide us with a general method of actually finding out whether a given sentence belongs to the set; for in general there is no effective decision procedure which, for any given sentence S, determines in a finite number of steps whether S is deducible from the axioms. But in any event, the standard construal assumes axiomatization only for the formulas of the uninterpreted calculus C rather than for all the sentences asserted by T,8 so that the proposed supporting argument does not actually apply here.

One of the attractions the standard construal has had for philosophers lies no doubt in its apparent ability to offer neat solutions to philosophical problems concerning the meaning and the reference of theoretical expressions. If the characteristic vocabulary of a theory represents “new” concepts, not previously employed, and designed specifically to describe the theoretical scenario, then it seems reasonable, and indeed philosophically important, to inquire how their meanings are specified. For if they should have no clearly determined meanings, then, it seems, neither do the theoretical principles in which they are invoked; and in that case, it would make no sense to ask whether those principles are true or false, whether events of the sort called for by the theoretical scenario do actually occur, and so forth.9 The answer that the standard construal is often taken to offer is, broadly speaking, that the meanings of theoretical terms are determined in part by the postulates of the calculus, which serve as “implicit definitions” for them; and in part by the correspondence rules, which provide them with empirical content. But this conception is open to various questions, some of which will be raised as we proceed.

As for the merits of axiomatization, its enormous significance for logic and mathematics and their metatheories needs no acknowledgment here. In some instances, axiomatic studies have served also to shed light on philosophical problems concerning theories in empirical science. One interesting example is Reichenbach’s axiomatically oriented, though not strictly formalized, analysis of the basis and structure of the theory of relativity.10 This analysis, which was undertaken some forty years ago, is technically distinct from more recent rigorous axiomatic formalization.

---


9 Various facets of this problem are carefully presented and explored in Nagel, The Structure of Science, chapter 6, and in Scheffler, The Anatomy of Inquiry, part II.

10 See H. Reichenbach, Axiomatik der relativistischen Raum-Zeit-Lehre (Braunschweig: Friedrich Vieweg und Sohn, 1924); and also Reichenbach’s article, “Ueber die physikalischen Konsequenzen der relativistischen Axiomatik,” Zeitschrift für Physik, 34 (1925), 32–48. In this article, to which Professor A. Grünbaum kindly called my attention, Reichenbach sets forth the main objectives of his axiomatic
Carl G. Hempel

zations; but it was nonetheless philosophically stimulating and illuminating, for it sought to clarify—much in the spirit of Einstein, I think—the roles of experience and convention in physical theorizing about space, time, and motion and the physical basis of the relativistic theory of spatial and temporal distances, of simultaneity, and so forth. More fundamentally, Reichenbach’s investigations were intended as a critique, based on a specific case study, of the Kantian notion of a priori knowledge. Again, an axiomatic approach played an important role in von Neumann’s argument that it is impossible to supplement the formalism of quantum mechanics by the introduction by hidden parameters in a way that yields a deterministic theory.

Some contemporary logicians and philosophers of science consider the axiomatization of scientific theories so important for the purposes of both science and philosophy that they have expended much effort and shown remarkable ingenuity in actually constructing such axiomatic formulations. Some of these, such as those developed by Kyburg, are small and relatively simple fragments of scientific theories in first-order logic; others, especially those constructed by Suppes and his associates, deal with richer, quantitative theories and formalize these with the more powerful apparatus of set theory and mathematical analysis.

But some of the claims that have been made in support of axiomatizing scientific theories are, I think, open to question. For example, Suppes has argued that formalizing and axiomatizing scientific concepts and theories is a “primary method of philosophical analysis,” and thus helps to “clarify conceptual problems and to make explicit the foundational assumptions of each scientific discipline,” and that to “formalize a connected family of concepts is one way of bringing out their meaning in an explicit fashion.”

In what sense can an uninterpreted axiomatization be said to “bring out the meanings” of the primitive terms? The postulates of a formalized efforts; on pp. 37–38, he rejects as irrelevant to his enterprise Hermann Weyl’s objection that Reichenbach’s axiomatization is too complicated and opaque from a purely mathematical point of view.


“STANDARD CONCEPTION” OF SCIENTIFIC THEORIES

theory are often said to constitute “implicit definitions” of the primitives, requiring the latter to stand for kinds of entities and relations which jointly satisfy the postulates. If axiomatization is to be viewed as somehow defining the primitives, then it is logically more satisfactory to construe axiomatization, with Suppes, as yielding an explicit definition of a higher order set-theoretical predicate. In either case, the formalized theory is then viewed in effect as dealing with just such kinds of entities and relations as make the postulates true.

This construal may have some plausibility for axiomatized purely mathematical theories—Hilbert adopted it in regard to his axiomatization of Euclidean geometry—but it is not plausible at all to hold that the primitive terms of an axiomatized theory in empirical science must be understood to stand for entities and attributes of which the postulates, and hence also the theorems, are true; for on this construal, the truth of the axiomatized theory would be guaranteed a priori, without any need for empirical study.

There are indeed cases in which axiomatization may be said to have contributed very significantly to the analytic clarification of a system of concepts. Suppes rightly mentions Kolmogorov’s axiomatization of probability theory as an outstanding example. But it should be noted that Kolmogorov’s formal system admits of such diverse interpretations as Carnap’s logical or inductive probability, Savage’s personal probability, and the empirical construal of probability in terms of long-run relative frequencies. The latter, of central importance in empirical science, has presented vexing difficulties to philosophical efforts at a satisfactory explication. Von Mises, Reichenbach, Popper, Braithwaite, and others all have sought to explicate the concept of statistical probability, or to specify the

15. For a careful and illuminating critical examination of the construal of postulates as implicit definitions for the primitives see chapter II of R. Grandy, “On Formalization and Formalistic Philosophies of Mathematics” (doctoral dissertation, Princeton University, 1967). Concerning the restrictions that the requirement of truth for the postulates imposes on the permissible interpretations of the primitives, Grandy notes that “it is not a restriction on the constants alone but on the set of constants plus the universe of discourse. A paraphrase of this is: The postulates implicitly define, if anything, the constants plus the quantifiers” (p. 41).

16. Kyburg therefore divides the axioms of a theory into “material axioms” and meaning postulates (in the sense of Carnap and Kemeny) and stresses that “we cannot lump [these] together and regard them as an implicit definition of the terms that occur in them” (Philosophy of Science, p. 124). It is presumably the meaning postulates alone that provide implicit definitions; but the distinction of two kinds of axioms is beset by the same difficulties as the analytic-synthetic distinction.

principles governing its scientific use. Some of these principles concern the pure calculus of probability, with which alone Kolmogorov’s axiomatization is concerned; others—and indeed the philosophically most perplexing ones—concern its application. And Kolmogorov’s analysis does not touch at all on this second part of the problem of “bringing out the meaning” of the term ‘probability’ “in an explicit fashion.”

Generally speaking, the formalization of the internal principles as a calculus sheds no light on what in the standard construal is viewed as its interpretation; it sheds light at best on part of the scientific theory in question. And as for the claim that formalization makes explicit the foundational assumptions of the scientific discipline concerned, it should be borne in mind that axiomatization is basically an expository device, determining a set of sentences and exhibiting their logical relationships, but not their epistemic grounds and connections. A scientific theory admits of many different axiomatizations, and the postulates chosen in a particular one need not, therefore, correspond to what in some more substantial sense might count as the basic assumptions of the theory; nor need the terms chosen as primitive in a given axiomatization represent what on epistemological or other grounds might qualify as the basic concepts of the theory; nor need the formal definitions of other theoretical terms by means of the chosen primitives correspond to statements which in science would be regarded as definitively true and thus analytic. In an axiomatization of Newtonian mechanics, the second law of motion can be given the status of a definition, a postulate, or a theorem, as one pleases; but the role it is thus assigned within the axiomatized system does not indicate whether in its scientific use it functions as a definitional truth, as a basic theoretical law, or as a derivative one (if indeed it may be said to have just one of these functions).

Hence, whatever philosophical illumination may be obtainable by presenting a theory in axiomatized form will come only from axiomatization of some particular and appropriate kind rather than just any axiomatization or even a formally especially economic and elegant one.

Suppes himself acknowledges that the “difficulty with the purely set-theoretical characterization of Kolmogorov is that the concept of probability is not sufficiently categorical” (ibid.), and he stresses that the interpretation of a formalized theory is logically much more complex than the talk of correspondence rules in “the standard sketch of scientific theories” would suggest (P. Suppes, “What Is a Scientific Theory?” in S. Morgenbesser, ed., Philosophy of Science Today, New York and London: Basic Books, 1967, pp. 55–67).

4. The Role of Pre-Theoretical Concepts in Internal Principles

The assumption, in the standard construal, of an axiomatized uninterpreted calculus as a constituent of a theory seems to me, moreover, to obscure certain important characteristics shared by many scientific theories. For that assumption suggests that the basic principles of a theory—those corresponding to the calculus—are formulated exclusively by means of a “new” theoretical vocabulary, whose terms would be replaced by variables or by dummy constants in the axiomatized calculus C. In this case, the conjunction of the postulates of C would be an expression of the type $\psi(t_1, t_2, \ldots, t_n)$, formed from the theoretical terms by means of logical symbols alone. Actually, however, the internal principles of most scientific theories employ not only “new” theoretical concepts but also “old,” or pre-theoretical, ones that are characterized in terms of the antecedent vocabulary. For the theoretical scenario is normally described in part by means of terms that have a use, and are understood, prior to, and independently of, the introduction of the theory. For example, the basic assumptions of the classical kinetic theory of gases attribute to atoms and molecules such characteristics as masses, volumes, velocities, momenta, and kinetic energies, which have been dealt with already in the antecedent study of macroscopic objects; the wave theory of light uses such antecedently available concepts as those of wavelength and wave frequency; and so forth. Thus, the internal principles of a theory—and hence also the corresponding calculus C—have to be viewed, in general, as containing pre-theoretical terms in addition to those of the theoretical vocabulary. Accordingly, the conjoined postulates of C would form an expression of the type $\psi(t_1, t_2, \ldots, t_n, p_1, p_2, \ldots, p_m)$, where the t’s again correspond to “new” theoretical terms, while the p’s are pre-theoretical, previously understood ones. Consequently, the theoretical calculus that the standard conception associates with a theory is not, as a rule, a totally uninterpreted system containing, apart from logical and mathematical symbols, only new theoretical terms.

It might be objected, from the vantage point of a narrow operationalism, that in this new context, the “old” terms $p_1, p_2, \ldots, p_m$ represent new concepts, quite different from those they signify in their pre-theoretical employment. For the use of such terms as ‘mass,’ ‘velocity,’ and ‘energy’ in reference to atoms or subatomic particles requires entirely new operational criteria of application, since at the atomic and subatomic levels the
quantities in question cannot be measured by means of scales, electrometers, and the like, which afford operational criteria for their measurement at the pre-theoretical level of macroscopic objects. On the strict operationist maxim that different criteria of application determine different concepts, we would thus have to conclude that, when used in internal principles, the terms \( p_1, p_2, \ldots, p_m \) stand for new concepts, and that it is therefore improper to use the old pre-theoretical terms in theoretical contexts: that they should be replaced here by appropriate new terms, which, along with \( t_1, t_2, \ldots, t_k \), would then belong to the theoretical vocabulary.

But differences in operational criteria of application, as is well known, cannot generally be regarded as indicative of differences in the concepts concerned; otherwise, it would have to be held impossible to measure “one and the same quantity” in a particular instance—such as the temperature or the density of a given body of gas—by different methods, or even with different instruments of like construction; as a consequence, the diversity of methods of measuring a quantity, already at the macroscopic level, would call for a self-defeating endless proliferation and distinction of concepts of temperature, of concepts of density, and so forth.

Moreover, as long as we allow ourselves to use the notoriously vague and elusive notion of meaning, we will have to regard the meanings of scientific terms as reflected not only in their operational criteria of application, but also in some of the laws or theoretical principles in which they function. And in this context, it seems significant to note that some of the most basic principles that govern the pre-theoretical use (relative to the classical kinetic theory, let us say) of such terms as ‘mass,’ ‘velocity,’ and ‘energy’ are carried over into their theoretical use. Thus, in the classical kinetic theory, mass is taken to be additive in the sense that the mass of several particles taken jointly equals the sum of the masses of the constituents, exactly as for macroscopic bodies. Similarly, the conservation laws for mass, energy, and momentum and the laws of motion are—at least initially—carried over from the pre-theoretical to the theoretical level.

In fact, the principle of additivity of mass is here used not only as a pre-theoretical and as an internal theoretical principle, but also as a bridge principle. In the latter role, it implies, for example, that the mass of a body of gas equals the sum of the masses of its constituent molecules; it thus connects certain features of the theoretical scenario with corresponding features of macroscopic systems that can be described in pre-theoretical terms. Those different roles of the additivity principle are clearly presupposed in the explanation of the laws of constant and of multiple proportions, and in certain methods of determining Avogadro’s number. These considerations suggest that the term ‘mass’ and others can hardly be taken to stand for quite different concepts, depending on whether they are applied to macroscopic objects or to atoms and molecules.

In support of the same point, it might be argued also that classical mechanics imposes no lower bounds on the size of the mass of the bodies to which the concepts of mass, velocity, kinetic energy, etc., can be significantly applied, and the laws governing these concepts are subject to no such restrictions either. This suggests a further response to the operationist objection considered a moment ago: the application of classical mechanical principles indicates that macroscopic methods using mechanical precision scales, etc., are not sufficiently sensitive for weighing atoms, but that certain indirect procedures will provide operational means for determining their masses. Accordingly, the need for different methods of measurement indicates, not a conceptual difference in the meanings of the word ‘mass’ as used in the two contexts, but a large substantive difference in mass between the objects concerned.

Analogous arguments, however, are not applicable in every case where pre-theoretical terms are used in the formulation of theoretical principles. According to current theory, for example, the mass of an atomic nucleus is less than the sum of the masses of its constituent protons and neutrons; thus the principles of additivity—and of conservation—of mass are abandoned at the subatomic level. Are we to say that this “theoretical change” indicates a change in the meaning of the term ‘mass,’ or rather that there has been a change in certain previously well-entrenched general laws which, before the advent of the new theory, had been erroneously believed to hold true of that one quantity, mass, to which both the new theory and the earlier one refer?

This question has received much attention in recent years in the debate over the ideas of Feyerabend, Kuhn, and some others concerning theoretical change in science and the theory-dependence of the meanings of scientific terms. As the debate has shown, however, a satisfactory reso-

---

19 This point is made also by Achinstein, Concepts of Science, p. 114; indeed his discussion, on pp. 106–119, of ways in which theoretical terms are introduced in science presents many illuminating observations and illustrations which accord well with the view expressed in this section, and which lend further support to it.

20 See, for example, T. S. Kuhn, The Structure of Scientific Revolutions (Chicago:
Carl G. Hempel

lution of the issue would require a more adequate theory of the notion of sameness of meaning than seems yet to be at hand.

5. The Role of a Model in the Specification of a Theory

As mentioned earlier, some adherents of the standard construal regard a theory as having a third component, in addition to the calculus and the rules of correspondence, namely, "a model for the abstract calculus, which supplies some flesh for the skeletal structure in terms of more or less familiar conceptual or visualizable materials." 21

In Bohr's theory of the hydrogen atom, for example, the postulates of the calculus would be the basic mathematical equations of the theory, expressed in terms of uninterpreted symbols such as 'i', 'r', 'E_i.' The model specifies the conception, referred to earlier, of a hydrogen atom as consisting of a nucleus circled by an electron in one or another of the orbits available to it, etc. In this model, 'r' is interpreted as the radius of the ith orbit, 'E_i' as the energy of the electron when in the ith orbit, etc. The correspondence rules, finally, link the theoretical notion of energy emission associated with an orbital jump to the experimental concept of corresponding wavelengths or spectral lines, and they establish other linkages of this kind.

In discussing these three components of Bohr's theory, Nagel remarks that as a rule the theory is embedded in a model rather than being formulated simply as an abstract calculus and a set of correspondence rules because, among other reasons, the theory can then be understood with greater ease than the inevitably more complex formal exposition. 22 It seems, however, that in some cases the significance of models in Nagel's sense goes further than this, as I will try to indicate briefly.

The term 'model' has been used in several different senses in the philosophy of science. One of these pertains to what might be called analogical models, such as the mechanical or hydrodynamic representations of electric currents or of the luminiferous ether that played a considerable role in the physics of the late nineteenth and early twentieth centuries. Models of this kind clearly are not intended to represent the actual microstructure of the modeled phenomena. They carry an implicit 'as if' clause with them; thus, electric currents behave in certain respects as if they consisted in the flow of a liquid through pipes of various widths and under various pressures; the analogy lies in the fact that phenomena of the two different kinds are governed by certain laws that have the same mathematical form. Analogical models may be of considerable heuristic value; they may make it easier to grasp a new theory, and they may suggest possible implications and even promising extensions of it; but they add nothing to the content of the theory and are, thus, logically dispensable.

But this verdict does not seem to me to apply to what Nagel would call the models implicit in such theories as the kinetic theory of gases, the classical wave and particle theories of light, Bohr's theory of the hydrogen atom, the molecular-lattice theory of crystal structure, or recent theories of the molecular structure of genes and the basis of the genetic code. All these claim to offer, not analogies, but tentative descriptions of the actual microstructure of the objects and processes under study. Gases are claimed actually to consist of molecules moving about and colliding at various high speeds, atoms are claimed to have certain subatomic constituents, and so forth. To be sure, these claims, like those of any other scientific hypothesis, may subsequently be modified or discarded; but they form an integral part of the theory. For example, as I suggested earlier, if a model in Nagel's sense characterizes certain theoretical variables as masses, velocities, energies, and the like, this may be taken to indicate that certain laws which are characteristic of masses, velocities, and energies apply to those variables, and that, if some of those laws are suspended in the theory, the requisite modifications will be made explicit. This happened, for example, in Bohr's model, where—in contrast to classical electromagnetic theory—an orbiting electron is assumed to radiate no energy. Hence, the specification of the model determines in part what consequences may be derived from the theory and, hence, what the theory can explain or predict.
More specifically, it seems that when a scientific theory is axiomatized, the process is limited to the mathematical connections that the theory assumes between quantitative features of the scenario; other theoretically relevant aspects of the scenario are specified by means of a model. I therefore agree with Sellars who remarks in a very similar vein that "in actual practice . . . the conceptual texture of theoretical terms in scientific use is far richer and more finely grained than the texture generated by the explicitly listed postulates," and that, in particular, the "thingish or quasi-thingsh character of theoretical objects, their conditions of identity . . . are some of the more familiar categorial features conveyed by the use of models and analogies." Thus, a model in the sense here considered is not only of didactic and heuristic value: The statements specifying the model seem to me to form part of the internal principles of a theory and as such to play a systematic role in its formulation.

It must be acknowledged that this way of formulating part of the internal principles of a theory is not fully specific and precise, that it does not provide an unequivocal characterization of exactly what statements the theory is meant to assert. But axiomatization, in the form of a "calculus," of part of a theory does not satisfy this desideratum, either; for it does not cover the correspondence rules; and for these, too, it seems virtually impossible to provide a formulation that could be regarded as adequate and complete. Indeed, as Nagel remarks, "theories in the sciences . . . are generally formulated with painstaking care and . . . the relations of theoretical notions to each other . . . are stated with great precision. Such care and precision are essential if the deductive consequences of theoretical assumptions are to be rigorously explored. On the other hand, rules of correspondence for connecting theoretical with experimental ideas generally receive no explicit formulation; and in actual practice the coordinations are comparatively loose and imprecise."

6. The Status of Correspondence Rules

In the standard construal, schematized by (2) above, R is conceived as a class of sentences that assign empirical content to the expressions of the calculus; and their designation as operational definitions, coordinative

 definitions, or rules of correspondence conveys the suggestion that they have the status of metalinguistic principles which render certain sentences true by terminological convention or legislation. The sentences thus declared true—let us call them interpretative sentences—would belong to an object language containing both the calculus and the pre-theoretical terms employed in its interpretation. The theoretical terms in the calculus are then best thought of as "new" constants that are being introduced into the object language by means of the correspondence rules for the purpose of formulating the theory. The interpretative sentences might have the form of explicit definition sentences (biconditionals or identities) for theoretical terms, or they might be of a more general type, providing only a partial specification of meaning for theoretical sentences, perhaps in the manner of Carnap's reduction sentences or still more flexible devices. But at any rate they would be sentences whose truth is guaranteed by the correspondence rules.

But such a conception of correspondence rules is untenable for several reasons, among them the following:

First, scientific statements that are initially introduced by "operational definitions" or more general rules of application for scientific terms—such as the statements characterizing length by reference to measurement with a standard rod, or temperature in terms of thermometer readings—usually change their status in response to new empirical findings and theoretical developments. They come to be regarded as statements which are simply false in their original generality, though perhaps very nearly true within a restricted range of application, and possibly only under additional precautionary conditions. Most sentences warranted by operational definitions or criteria of application are eventually qualified as, strictly speaking, false by the very theories in whose development they played a significant role. Much the same point is illustrated by the following example: To "define" in experimental terms equal intervals of time, some periodic process may be chosen to serve as a standard clock, such as the swinging of a pendulum or the axial rotation of the earth as reflected in the periodic, apparent daily motion of some fixed star. The time intervals marked off by the chosen process are then equal by convention or stipulation. Yet it may happen that certain laws or theoretical principles originally based on evidence that includes the readings of standard clocks give rise to the

---


---

25 Such, perhaps, as interpretative systems of the kind I suggested in section 8 of "The Theoretician's Dilemma."
verdict that those clocks do not mark off strictly equal time intervals. One
striking example is the use of ancient astronomical reports of an almost
purely qualitative character—concerning the date and very roughly the
time of day when a certain total solar eclipse was observed at a given place
—to establish a very slow deceleration of the earth’s axial rotation, with a
consequent slow lengthening of the mean solar day (by no more than .003
seconds in a century).26

Thus, even though a sentence may originally be introduced as true by
stipulation, it soon joins the club of all other member statements of the
theory and becomes subject to revision in response to further empirical
findings and theoretical developments. As Quine has said, “convention-
ality is a passing trait, significant at the moving front of science, but useless
in classifying the sentences behind the lines.”27

These considerations might invite the following reply: Of course a
theory—including its correspondence rules—may well undergo changes in
response to new empirical findings; the question at issue, however, does
not concern the possible effects of scientific change on correspondence
rules, but rather the epistemic status of the interpretative sentences of a
given theory, “frozen,” as it were, at a particular point of its development.
If such a theory is systematically characterized by means of a calculus and
a set of interpretative sentences, do not the latter have the character of
terminological conventions?

Here it should be recalled, first of all, that a theory usually links a given
theoretical concept to several distinct kinds of phenomena that are charac-
terizable in terms of the antecedently available vocabulary. For example,
contemporary physical theory provides for several different ways of de-
termining Avogadro’s number or the charge of an electron or the velocity
of light. But not all the interpretative sentences thus provided for a given
theoretical term can be true by convention; for they imply statements to
the effect that if one of the specified methods yields a certain numerical
value for the quantity in question, then the alternative methods will yield
the same value, and whether this is in fact the case is surely an empirical
matter and cannot be settled by terminological convention. This point has
indeed been stressed by some proponents of the standard conception.

Cf. R. Carnap, “Testability and Meaning,” Philosophy of Science, 3 (1936),
comment applies to interpretative systems, of course; see my “The Theoretician’s Dilem-
ma,” p. 74.

Perhaps Quine’s earliest detailed attack on the distinction is mounted in his
classical “Two Dogmas of Empiricism” (1951), reprinted in W. V. O. Quine, From
a Logical Point of View, 2nd ed. (Cambridge, Mass.: Harvard University Press,
1961). Another early critique is given in M. G. White, “The Analytic and the Syn-
7. On “Specifying the Meanings” of Theoretical Terms

Our critical scrutiny, however, has suggested no solution to one central question which the standard construal sought to answer, namely the question of how the meanings of the “new” terms in a theory are specified. We found difficulties both with the conception that the postulates of the uninterpreted calculus provide implicit definitions for the theoretical terms and with the idea of correspondence rules as principles of empirical interpretation; but no alternative answer to the question has been offered. I believe now that the presumptive problem “does not exist,” as Putnam has said and argued, or, as I would rather say, that it is misconceived. In conclusion, I will briefly suggest some considerations in support of this view.

What reasons are there for thinking that the “new” concepts introduced by a theory are—or at least should be—specifiable by means of the antecedently available vocabulary? One consideration that influenced my earlier concern with the problem is, briefly, to this effect: A theory purports to describe certain facts, to make assertions that are either true or false. But a sentence will qualify for the status of being either true or false only if the meanings of its constituent terms are fully determined; and if we want to understand a theory, or to examine the truth of its claims, or to apply it to particular situations, we must understand the relevant terms, we must know their meanings. Thus, an adequate statement of a theory will require a specification of the meanings of its terms—and what other means is there for such a specification than the antecedently available vocabulary?

But even if, for the sake of argument, we waive questions about the concept of meaning here invoked, these considerations are not compelling. On the contrary; when at some stage in the development of a scientific discipline a new theory is proposed, offering a changed perspective on the subject matter under study, it seems highly plausible that new concepts will be needed for the purpose, concepts not fully characterizable by means of those antecedently available. This view seems to me to derive support from those studies of the language of science—especially in the logical empiricist tradition—which have led to a steady retrenchment of the initial belief in, or demand for, full definability of all scientific terms by means of some antecedent vocabulary consisting of observational predicates or the like. The reasons that led to countenancing the introduction

of new terms by means of reduction sentences, interpretative systems, or probabilistic criteria of application all support the idea that the concepts used in a new scientific theory cannot be expected always to be fully characterizable by antecedently available ones.

But the very relaxation of the requirements for the introduction of new scientific terms gave rise to such questions as whether we can claim to understand such partially interpreted terms; whether the sentences containing them can count as significant assertions or can be regarded at best as an effective, but inherently meaningless, machinery for inferring significant statements, couched in fully understood terms, from other such statements; and whether reliance on incompletely interpreted theoretical terms could be entirely avoided in science.

But this way of looking at the issue presupposes that we cannot come to understand new theoretical terms except by way of sentences specifying their meanings with the help of previously understood terms; and surely this notion is untenable. We come to understand new terms, we learn how to use them properly, in many ways besides definition: from instances of their use in particular contexts, from paraphrases that can make no claim to being definitions, and so forth. The internal principles and bridge principles of a theory, apart from systematically characterizing its content, no doubt offer the learner the most important access to an “understanding” of its expressions, including terms as well as sentences.

To be sure, all these devices still leave unanswered various questions concerning the proper use of the expressions in question; and this may seem to show that, after all, the meanings of those expressions have not been fully specified, and that the expressions therefore are not fully understood. But the notion of an expression that has a fully specified meaning or an expression that is fully understood is obscure; besides, even for terms that are generally regarded as quite well understood there are open questions concerning their proper use. For example, there are no sharp criteria that would determine, for any strange object an astronaut might encounter on another planet, or indeed for any object that might be produced in a test tube on earth, whether it counts as a living organism. Theoretical concepts, just like the concept of living organism, are “open-ended”; but that, evidently, is no bar to their being adequately understood for the purposes of science.