theoretical terms are introduced, and come to be understood, much in the way most words are introduced and understood; there is no reason to believe that the process can always be explicated or rationally reconstructed in accordance with the presupposition just mentiond.

'introduced and come to be understood' – yes – but that is not the most salient point for understanding their role and meaning in a theory; what eventually will – should – be made precise in some more formal and b[illegible], explicative sense

Discussion

Professor Suppes:

It is a pleasure to engage in discussion again with Professor Hempel. We had a debate on some of these topics a few months ago.[1]

I would like to comment briefly on some of the issues he raises. I think they are difficult issues to sharpen, but I will not spend long since it is not my purpose to reach for equal time. In dealing with these issues I would like to try to bring one or two specific arguments to bear on the usefulness of formalization in science. Rather than deal with the broad issues, let me take a try at one or two very specific ones. First, a passing remark on the matter of implicit definitions and the meaning of concepts. The matter was probably first made explicit in a classical exchange between Frege and Hilbert. The purpose of formalization is primarily to give an explicit definition of the overall concept, not an implicit definition of the subsidiary concepts.

Second, there are issues in the philosophy of science and also of scientific interest that require formalization for settling or clarifying the issues. Let me mention two rather distinct matters that are very lively issues among some philosophers of science and some scientists. First, a very large body of current literature on the status of probability in quantum mechanics has been developed by physicists, philosophers, and mathematicians. Without formalization it simply is impossible to resolve the issues that have now become very complex regarding the role and the position of classical probability in quantum mechanics. We cannot make any further progress in pinning down the complex set of issues that are now in front of us simply by informal remarks and by intuitive interpretations. Also related to these problems is the argument about the status of classical and nonclassical logic in quantum mechanics; and here again I think we are beyond the stage at which further progress can be made by any other means than very careful, explicit, and adequate formalization.

[1] At the Eastern Division Meetings of the American Philosophical Association, December, 1968, where Hempel served as a commentator on Suppes [1968].

The second example, which is from a different domain but shows that formalization has the same role to play, is the issue between the behavioral psychologists on the one hand and cognitive and linguistically oriented psychologists on the other regarding whether behavioral psychologists' concepts are rich enough to account for language learning and for language behavior. A great many words have been said and a great many debates[2] have taken place about this, and it seems to me that further clarification of the substantive scientific issue, which is also an interesting philosophical issue, will depend partly upon formalization. In this case clarification may very well depend ultimately upon further empirical work as well, but at the very least progress can be made on certain parts of the argument only with further formalization.

Finally, regarding the view that formalization is restricted primarily to the internal principles (what Professor Hempel calls the set I of principles) I would claim that, from the standpoint of scientific methodology, this very much is *not* the case. One of the most elaborate, sophisticated, and extensive efforts at formalization in the methodology of science takes place in the context of bridging rules—the entire effort of the last thirty years in mathematical statistics can be thought of as an attempt to put in very explicit and very formal terms the problems of methodology that are involved in the testing of hypotheses, the designing of experiments, and all the other similar questions that we are concerned with in any kind of complicated science. The thing that is very characteristic of this statistical literature is that it has been written in a very firm and very definite atmosphere of formalization exactly as a part of contemporary mathematics.

Professor Hempel:

Professor Suppes's first remarks touched on an interesting difficulty in the conception of a postulate system as providing implicit definitions for the primitives: by these "definitions," all of the primitives are characterized jointly, so that, in a manner of speaking, each primitive is "defined" in terms of the others. Thus, what meanings or what extensions may be assigned to one of the primitives always depends on what interpretations are given to the others. It may be preferable, therefore, to say that the constraints which the postulates impose on permissible interpretations concern, not the primitives individually, but the entire ordered set of them. This idea is implemented by a mode of axiomatization that Professor Suppes and his associates have ingeniously

[2] See, for example, Skinner [1957] and Chomsky's [1959] review of it.

applied to a variety of theories, namely, axiomatization by explicit definition of one set-theoretical predicate.[3] In the case of an axiomatized theory with n primitives, the predicate is defined so as to apply to an n-tuple of set-theoretical entities (sets, functions, and so on) just in case the interpretation of the primitives by those set-theoretical entities turns the postulates into true sentences. This procedure has certain logical attractions, but these do not invalidate any of the reservations I expressed concerning some of Professor Suppes's arguments for the desirability of axiomatization in science, for those arguments were essentially independent of the particular mode of axiomatic formalization. Let me acknowledge again, however, that axiomatization of theories can be of value for certain philosophical or scientific purposes; Professor Suppes has mentioned some pertinent examples in his remarks today.

As for axiomatizing the sentences that provide the interpretation of a formalized theoretical system, this is certainly possible, and in some cases, as in the theory of measurement, has yielded illuminating results. But the procedure again eventuates in a formal system, which must be given a specific interpretation if the intended specific content of the theory is to be captured. Thus, at some point, there will have to be an interpretation in terms already understood that does not in turn rely on formalization. Here is one context in our discussions where I think Professor Kuhn's paper is highly relevant, for it explores ways in which the requisite agreement in the understanding and the use of scientific terms may be attained by the members of a scientific community without reliance on, or even availability of, explicitly formulated criteria of application.

Professor Causey:

I would just like to ask Professor Hempel a couple of questions. The first is mainly a terminological one. If I am not mistaken you defined a theory, or at least the set of sentences of a theory, to be the set of all logical consequences of the union of the internal principles I and the bridge principles B. Although I do not think it is a terribly crucial matter, it seems to me that if we are going to talk about theories and use the term in any sense at all like it is used by working scientists, then this notion or this definition of a theory as a set of sentences is much too broad, because it will include in those sets of consequences all kinds of pretheoretical sentences which I would think the average scientist would not want to consider part of *his* theory.

[3] The method is explained and illustrated in Suppes [1957], Ch. 12.

Now a deeper question, and one which is rather general, is the following: You raised a number of good objections against the Received View. But so far as I understand your view of a theory in terms of internal principles *I* and bridge principles *B*, and to the extent that I understand at least most of your objections against the Received View, it seems to me that these objections also will apply to your own view —though perhaps in a slightly altered form.

Professor Hempel:

A scientist will formulate a theory by specifying only a finite subset, say *S*, of its intended assertions. He will not know all of the logical consequences of this set; and while affirming the theory, he might possibly doubt or even deny certain sentences which, without his knowing it, are consequences of *S*. Would this make it more appropriate to construe a theory as a set of sentences each of which its adherents would clearly feel themselves committed to assert?

For the purposes of a logical and methodological analysis, this would be unacceptable; for—to mention one reason—a careful experimental disconfirmation of certain logical consequences of *S* would count as disconfirming the theory even if some of its adherents did not feel themselves committed to give assent to those consequences. Hence, if a theory is at all represented by a set of sentences, then it must be construed as containing all consequences of any of its subsets. This in turn makes it necessary to reckon among the assertions of a theory also such consequences of *S* as would generally be regarded uninteresting or trivial, as well as those containing only antecedent terms. That a theory should yield consequences of this latter kind is, in fact, an essential requirement if theories are assumed to yield predictions couched in the antecedent vocabulary by deduction from sentences describing prior occurrences in antecedently available terms.

As for Professor Causey's second point, I do not think the conception of a theory as represented by internal principles and bridge principles, as in schema I, is open to essentially the same objections as the standard construal represented by schema II. In particular, the internal principles are not claimed to provide implicit definitions for the theoretical terms, nor are they assumed to be couched exclusively in the "new" theoretical vocabulary; similarly, the bridge principles are not regarded as assigning empirical content to the theoretical terms, nor are they held to be established by convention or by rules; they are taken to be statements on a par with the internal principles and subject to possible revocation if the theory should be found to conflict with well-established empirical findings.

However, as I acknowledged in my lecture,[4] I have offered no precise criteria for delimiting the two sets. In particular, the dividing line cannot be drawn by reference to the constituent nonlogical terms since both sets normally contain antecedent as well as theoretical terms; nor is there a distinguishing difference in epistemic status, such as acceptance by terminological convention *vs.* acceptance on empirical grounds. It is thus assuredly a vague distinction, but I hope it may be sufficiently clear and suggestive in the context in which I have used it, namely, as a background for a discussion of the standard conception.

Professor Achinstein:

You have given up one distinction and are advocating another one; you have given up the observational-theoretical distinction and you are advocating a distinction between theoretical terms and antecedently available terms, and now I am wondering whether this distinction does not introduce as many problems as the old one.

One preliminary point: you mentioned the Bohr theory as your first example; now each term that Bohr actually used was available prior to the Bohr theory. Some of the terms were used by Rutherford, and even before Rutherford, terms like "electron" and "nucleus." Indeed, I would venture the claim that with respect to most theories most of the terms in that theory that Positivists and others traditionally have called theoretical were available before that theory was formulated.

Second, when you speak of the "antecedently available vocabulary" do you have in mind simply words or do you mean the concepts behind them? For example, if in their theory of mechanics physicists had used the term "energy of motion" first and then someone came along and used the term "kinetic energy" instead with the same meaning, would the latter then be a theoretical term in the theory? One would be tempted to say here that no new concept has been introduced even though a new term has. On the other hand, Clausius introduced the term "entropy" which he defined explicitly by reference to the antecedently available terms "heat" and "temperature" and the notion of the mathematical integral. In this case one is tempted to say that a new concept has been introduced even though the term "entropy" is explicitly defined by means of the antecedently available vocabulary. So when you speak of the theoretical terms of the theory are you referring simply to the terms or to the concepts they express? If the latter, what criteria do you use to decide whether a concept is a new one or is antecedently available?

[4] See Hempel [1970], end of Sec. 6; the point is not mentioned in the summary-abstract above.

Professor Hempel:

The two vocabularies are assumed to consist of scientific terms in their specific scientific modes of use, or with their specific scientific meanings or interpretations. Even the representation, in the standard conception, of theoretical terms by uninterpreted constants is not meant, of course, to characterize those terms as devoid of meaning: it is an analytic device intended to explicate the logical means by which, and the extent to which, the theory may be said to specify the meanings—or the extensions—of the theoretical terms within a conceptual and linguistic framework antecedently available. The two principal logical devices envisaged by the standard construal, as I have mentioned, are implicit definition by theoretical postulates and empirical interpretation by rules of correspondence.

The standard construal often assumes that the empirical interpretation is effected by means of an observational vocabulary. To avoid the difficulties of this latter notion and of the construal of all theoretical terms as nonobservational, I suggested that the basis for the empirical interpretation could more simply be taken to be a set of previously understood terms, no matter whether they stand for items which, in some plausible sense, might qualify as observable.

Professor Achinstein is right in what I take him to be suggesting, namely, that in the sense here intended, the theoretical terms of a given theory cannot be characterized simply as those extralogical words or symbols which occur in the formulation of the theory, but not in the antecedent vocabulary. For an old term might acquire a new interpretation when used in the theory; and in this case, the question of how its new meaning or its new extension is specified arises for the old terms no less than for the new ones. The sorting out of the theoretical terms would therefore require criteria determining whether a given antecedently available term as used in a theory has the same meaning—or the same reference—as in its pretheoretical use.

I do not know of any satisfactory explication of sameness of meaning, and I agree with Quine and other critics in regarding the idea as intrinsically unclear. But even if we ask only about changes of extension, I see considerable difficulties, for it is not clear how the extension of a term as used in a given theory is to be characterized in a nontrivial way.[5]

But the principal purpose of my paper was to raise certain questions about the standard construal of theories; and the questions I did

[5] Professor Hempel's reply to Professor Putnam, below, contains a slight amplification of this remark.

raise do not depend on exactly where the dividing line between the antecedent and the theoretical vocabulary is drawn.

Professor Bromberger:

I would like to ask a very simple-minded question: Exactly what is the status of this analysis of a theory? What function does it fulfill and how does it clarify? As an analysis of the concept of a theory, that is, the concept that is embodied in our use of the word 'theory' in English or equivalent ones in other languages, I think it is demonstratively false. Now I take it that there is some kind of a program associated with this general sketch. The view that a theory consists of an uninterpreted calculus and rules of interpretation had one great virtue as a program, namely that it was a program which, if successful, would have made explicit the rules that govern or that ought to govern acceptance and rejection of theories. It failed as a program and therefore it failed to fulfill this task, which I think is an important task of the philosophy of science. Now do you envisage your new distinction as entailing a program that will show how the theories that we in fact have might ultimately be analyzed? And would such an analysis then make explicit why they are accepted, how they are accepted, and whether their acceptance is based on principles that are philosophically kosher?

Professor Hempel:

Professor Bromberger is right in stressing the programmatic side of the standard conception. One of its objectives was to explicate, and appraise from the point of view of an analytic-philosophical conscience, the principles governing concept formation in scientific theories. Another objective was similarly to exhibit and appraise the principles governing the testing of scientific theories. On the standard view, the testing of a theory depends essentially on the deductive connections which the theory is taken to establish between empirical sentences couched in observational—or, more generally—in antecedent terms.[6]

My paper was intended principally as a criticism of the basic assumptions by means of which the standard construal tackles its tasks; I did not put forward a properly developed alternative. Let me just add that the standard construal of the connection between new theoretical concepts and antecedent ones seems to me overly restrictive, and that it appears to be prompted by an assumption which

[6] For details, see, for example, Hempel [1958], pp. 46–48, 75–76.

impressed me as quite reasonable when I wrote "The Theoretician's Dilemma,"[7] but which I now regard as unwarranted, namely that, to be "philosophically kosher" as Professor Bromberger calls it, the introduction of theoretical terms must somehow provide an explanation, in formally reconstructible ways, of the new concepts by means of the antecedent ones. The considerations that militate against this view also throw into doubt the conception of theories as establishing deductive transitions among sentences couched in terms of the antecedent vocabulary. But these are issues about which I am just trying to form clearer ideas, and thus, to my regret, I cannot be more specific about them now.

Professor Putnam:

I agree with almost all of the things you said. However, there are still I think two points where I think our views are in disagreement and I would like to get your reaction to one of them. You said something to the effect that if a term introduced into a theory had no exact meaning then it makes no sense to ask whether it is true or false. Now I think that is wrong—I think first of all that very few terms have what one could call an exact meaning, and anyway I suspect that scientists are not interested in the exact meaning, in the linguistic sense, of the terms they use; they would be very bored. And I suspect that is one mistake that really mattered. If meaning is an issue then it hangs over all the way because I want to get into problems about concepts and how I would know it is the same concept. This all gets cut short if one says at the outset that the question is not whether the terms of a theory have meaning, or what they mean in the sense of linguistic usage or anything like that—or whether they refer to anything. Of course, Quine would say this because he is pessimistic about the notion of meaning. I am not as pessimistic, but I do think that the notion of meaning is fairly irrelevant to this part of philosophy of science, and that one can proceed in a purely extensionalistic way. That would be the first thing on which I want to get your reaction. Then I think we would see that many of the questions are somewhat easier; they are still hard but at least they are a little different. For example, I do not face the question, when should I say a term in an older theory is the same concept as the term I am now using. I do face the question, when should I say it refers to something that I now recognize. Should I say that Bohr's term 'electron' referred to what I called electrons? And there I would say "yes" because I advocate a

[7] Hempel [1958].

principle of charity with respect to such questions: a lot of his beliefs are still true judging from the standpoint of my present theory, so rather than say that the poor man was not talking about anything at all, that it was like witches or phlogiston, I charitably say "Yes, he was talking about electrons although some of his sentences were false." And here I would say the great insight for me contained in Tarski's theory of truth is that anyone who grants me this mild thing that I can say what Bohr's terms were referring to has granted me the concept of truth for all of Bohr's sentences. You know, Donald Davidson has recently been emphasizing that it is one of the main ways one can look at Tarski: if you give me the reference, I have got truth.[8]

Put in another way to anticipate future polemics, if you are not convinced by Kuhn's talk about scientists being in different worlds, if you do not buy that as far as reference is concerned, if it is not true enough to lead me to say that I do not know what Bohr was referring to, then you cannot buy it as far as truth is concerned either.

Professor Hempel:

I feel quite dubious about the concept of meaning—in fact, I share what Professor Putnam calls Quine's pessimism on that score. I spoke about the meanings of scientific terms principally in order to reflect one of the philosophical assumptions that seem to me to have informed the standard conception. This is the idea that (*i*) a philosophical analysis of scientific theories should seek to explicate, among other things, by what logical means theoretical terms are assigned—or may be conceived as being assigned—specific meanings, given the linguistic and terminological apparatus available prior to the construction of the theory; and that (*ii*) unless a philosophically satisfactory explication of those means can be given, the sentences of the theory, insofar as they contain theoretical terms, would have to be regarded as lacking clearly determined meanings and, consequently, as being neither true nor false.

Professor Putnam is right in saying that to ensure truth or falsity for those sentences, it would suffice that their constituent nonlogical terms each have a definite reference; that the question of meanings or intensions may be disregarded in this context. But this still leaves the question whether the formulation of a theory does assign determinate extensions to the theoretical terms, and if so, whether this assignment is achieved or achievable in philosophically satisfactory ways by

[8] See Davidson [1967].

means of the conceptual and linguistic apparatus of science available when the theory is being constructed.

The answer would clearly be in the affirmative if it could be shown that a theory implies, for each of its theoretical terms, a biconditional sentence linking the term to an expression containing only antecedent terms, for the theory would then assign to the theoretical terms the same extension as to that expression. But this condition is rarely, if ever, met; and the standard conception may be regarded as suggesting a more modest answer, to the effect that the extensions of the theoretical terms are fixed in part by the postulates of the calculus and in part by the rules of correspondence. And in so far as my critical observations concerning the standard view carry weight, they raise questions also about its adequacy for the extensional characterization of theoretical terms.

Let me add some further thoughts on Professor Putnam's suggestion. Even though, as Quine has often stressed, the theory of reference is in philosophically much more satisfactory shape than the theory of meaning, there are special problems concerning the reference of theoretical terms which the general theory of reference does not automatically answer. In fact, these problems exhibit distinct analogies to those concerning the "meanings" of theoretical terms. Take Putnam's example about the references of the term 'electron' as used, say, in Bohr's first theory of the hydrogen atom and the term 'electron' as used in contemporary physics; let us abbreviate these as 'e_B' and 'e_C', respectively. To say, with Putnam, that the two terms have the same reference comes to asserting that

$$(x)(e_B x \equiv e_C x).$$

But to prove this, or even to make it plausible, we would need specifications of the two extensions. What sentences of each of the relevant theories should count as determining those extensions? (This is strongly reminiscent of the question: what sentences of a given theory are to count as determining the meaning of one of its theoretical terms?) Surely not all the sentences of each theory, for the two theories may well be incompatible. Besides, by asserting that the two terms have the same reference, Professor Putnam presumably wants to say that the two theories refer to the same set of objects but make different—perhaps incompatible—assertions about them. Hence it seems that the two extensions would have to be specified by two logically compatible characterizations of electrons that are implied by, but not logically equivalent to, the respective theories. But how are these characterizations to be chosen?

And—supposing this question were answered satisfactorily—how is the coextensiveness of the two characterizations to be established? It would hardly amount to a purely logical truth: but then, are we to use certain principles of Bohr's theory or of the contemporary theory, or perhaps some combination of the two kinds, in arguing for coextensiveness?

Thus, even if we concentrate our analytic efforts exclusively on the extensions of theoretical terms and the truth of theoretical sentences, some difficult problems remain.