RUDOLF CARNAP

The Methodological Character of Theoretical Concepts

I. Our Problems

IN DISCUSSIONS on the methodology of science, it is customary and useful to divide the language of science into two parts, the observation language and the theoretical language. The observation language uses terms designating observable properties and relations for the description of observable things or events. The theoretical language, on the other hand, contains terms which may refer to unobservable events, unobservable aspects or features of events, e.g., to micro-particles like electrons or atoms, to the electromagnetic field or the gravitational field in physics, to drives and potentials of various kinds in psychology, etc. In this article I shall try to clarify the nature of the theoretical language and its relation to the observation language. The observation language will be briefly described in Section 2 of this paper. Then a more detailed account of the theoretical language and the connections between the two languages will be given in Sections III–V.

One of the main topics will be the problem of a criterion of significance for the theoretical language, i.e., exact conditions which terms and sentences of the theoretical language must fulfill in order to have a positive function for the explanation and prediction of observable events and thus to be acceptable as empirically meaningful. I shall leave aside the problem of a criterion of significance for the observation language, because there seem to be hardly any points of serious disagreement among philosophers today with respect to this problem, at least if the observation language is understood in the narrow sense indicated above. On the other hand, the problem for the theoretical language is a very serious one. There are not only disagreements with respect to the exact location of the boundary line between the meaningful and the meaningless, but some philosophers are doubtful about the very possibility of drawing any boundary line. It is true that empiricists today generally agree that certain criteria previously proposed were too narrow; for example, the requirement that all theoretical terms should be definable on the basis of those of the observation language and that all theoretical sentences should be translatable into the observation language. We are at present aware that these requirements are too strong because the rules connecting the two languages (which we shall call "rules of correspondence") can give only a partial interpretation for the theoretical language. From this fact, some philosophers draw the conclusion that, once the earlier criteria are liberalized, we shall find a continuous line from terms which are closely connected with observations, e.g., 'mass' and 'temperature,' through more remote terms like 'electromagnetic field' and 'psi-function' in physics, to those terms which have no specifiable connection with observable events, e.g., terms in speculative metaphysics; therefore, meaningfulness seems to them merely a matter of degree. This skeptical position is maintained also by some empiricists; Hempel, for instance, has given clear and forceful arguments for this view (see his articles, (14) and (15)). Although he still regards the basic idea of the empiricist meaning criterion as sound, he believes that deep-going modifications are necessary. First, the question of meaningfulness cannot, in his opinion, be raised for any single term or sentence but only for the whole system consisting of the theory, expressed in the theoretical language, and the correspondence rules. And secondly, even for this system as a whole, he thinks that no sharp distinction between meaningful and meaningless can be drawn; we may, at best, say something about its degree of confirmation on the basis of the available observational evidence, or about the degree of its explanatory or predictive power for observable events.

The skeptics do not, of course, deny that we can draw an exact boundary line if we want to. But they doubt whether any boundary line is an adequate explication of the distinction which empiricists had originally in mind. They believe that, if any boundary line is drawn, it will be more or less arbitrary; and, moreover, that it will turn out to be either too narrow or too wide. That it is too narrow means that some terms or sentences are excluded which are accepted by scientists as meaningful; that it is too wide means that some terms or sentences are included which scientifically thinking men would not accept as meaningful.

My attitude is more optimistic than that of the skeptics. I believe that, also in the theoretical language, it is possible to draw an adequate boundary line which separates the scientifically meaningful from the meaningless. I shall propose criteria of significance; the criterion for theoretical terms will be formulated in Section VI, and the question of its adequacy will be examined in Section VII; the criterion for theoretical sentences will be given in Section VIII.

Two alternative forms for the introduction of scientific concepts into our two-language system will be explained and their comparative usefulness examined (Sections IX and X). One kind is that of theoretical concepts introduced into the theoretical language by postulates. The other kind I call "disposition concepts." They may be introduced into an extended observation language. Concepts defined by so-called operational definitions and the so-called intervening variables belong to this kind. I shall try to show that the introduction in the form of theoretical concepts is a more useful method because it allows greater freedom in the choice of conceptual forms; moreover, it seems more in accord with the way the scientists actually use their concepts.

In the last section, I discuss briefly the possibilities and advantages of the use of theoretical concepts in psychology.

II. The Observation Language L_O

The total language of science, L, is considered as consisting of two parts, the observation language L_O and the theoretical language L_T. I shall here briefly indicate the nature of L_O; the later discussion will chiefly concern L_T and its relations to L_O. Without actually specifying it, we assume that the logical structure of L_O is given. This would include a specification of the primitive constants, divided into logical and descriptive (i.e., nonlogical) constants. Let the observational vocabulary V_O be the class of the descriptive constants of L_O. Further, for each language part the admitted types of variables are specified. In L_O, it may suffice to use only individual variables, with observable events (including thing-moments) taken as individuals. Then rules of formation, which specify the admitted forms of sentences, and rules of logical deduction are given.

Let us imagine that L_O is used by a certain language community as a means of communication, and that all sentences of L_O are understood by all members of the group in the same sense. Thus a complete interpretation of L_O is given.

The terms of V_O are predicates designating observable properties of events or things (e.g., "blue," "hot," "large," etc.) or observable relations between them (e.g., "x is warmer than y," "x is contiguous to y," etc.).

Some philosophers have proposed certain principles which restrict either the forms of expression or the procedures of deduction in "the language," in order to make sure that everything said in the language is completely meaningful. It seems to me that the justification of such requirements depends upon the purpose for which the language in question is used. Since L_O is intended for the description of observable events and therefore is meant to be completely interpreted, the requirements, or at least some of them, seem to have merit. Let us consider the most important requirements that have been proposed for some or any language L.

1. Requirement of *observability* for the primitive descriptive terms.
2. Requirements of various degrees of strictness for the nonprimitive descriptive terms.
   (a) Explicit *definability*.
   (b) *Reducibility* by conditional definitions (e.g., by reduction sentences as proposed in (5)).
3. Requirement of *nominalism:* the values of the variables must be concrete, observable entities (e.g., observable events, things, or thing-moments).
4. Requirement of *finitism*, in one of three forms of increasing strictness:
   (a) The rules of the language L do not state or imply that the basic domain (the range of values of the individual variables) is infinite. In technical terms, L has at least one finite model.
   (b) L has only finite models.
   (c) There is a finite number n such that no model contains more than n individuals.
5. Requirement of *constructivism:* every value of any variable of L is designated by an expression in L.
6. Requirement of *extensionality*. The language contains only truth-functional connectives, no terms for logical or causal modalities (necessity, possibility, etc.).

Any language fulfilling these requirements is more directly and more completely understandable than languages transgressing these limitations. However, for the language as a whole, the requirements are not justified; we shall reject them later for the theoretical language L_T.

Since then we have in the part L_T all the freedom of expression desired, we may well accept some or all of these requirements for L_O.

We have already accepted requirements 1 and 3. The decision about requirement 2 depends upon our intention concerning disposition terms (e.g., "soluble," "fragile," "flexible"). We shall not include them in L_O itself; thus L_O is here taken as a *restricted observation language* fulfilling the stronger requirement 2(a). Later (in Section IX) the possibility of an extended observation language L'_O, which allows the introduction of disposition terms, will be explained. Another method consists in representing the disposition concepts by theoretical terms in L_T (Section X).

The weakest requirement 4(a) of finitism is fulfilled in L_O. Therefore it is easily possible to satisfy requirement 5. Further, we take L_O as an extensional language; thus requirement 6 is fulfilled.

III. The Theoretical Language L_T

The primitive constants of L_T are, like those of L_O, divided into logical and descriptive constants. Let the theoretical vocabulary V_T be the class of the descriptive primitive constants of L_T. We shall often call these constants simply "theoretical terms." (They are often called "theoretical constructs" or "hypothetical constructs." However, since the term "construct" was originally used for explicitly defined terms or concepts, it might be preferable to avoid this term here and use instead the neutral phrase "theoretical term" (or "theoretical primitive"). This use seems to be in better accord with the fact that it is, in general, not possible to give explicit definitions for theoretical terms on the basis of L_O.)

We may take it for granted that L_T contains the usual truth-functional connectives (e.g., for negation and conjunction). Other connectives, e.g., signs for logical modalities (e.g., logical necessity and strict implication) and for causal modalities (e.g., causal necessity and causal implication) may be admitted if desired; but their inclusion would require a considerably more complicated set of rules of logical deduction (as syntactical or semantical rules). The most important remaining problem for the specification of the logical structure concerns the ranges of values for the variables to be admitted in universal and existential quantifiers, and thereby the kinds of entities dealt with in L_T. This problem will be discussed in Section IV.

A *theory* is given, consisting of a finite number of *postulates* formulated in L_T. Let T be the conjunction of these postulates. Finally, *correspondence rules* C are given, which connect the terms of V_T with those of V_O. These rules will be explained in Section V.

IV. The Problem of the Admissibility of Theoretical Entities

It seems that the acceptance of the following three conventions *C1*–*C3* is sufficient to make sure that L_T includes all of mathematics that is needed in science and also all kinds of entities that customarily occur in any branch of empirical science.

Conventions about the domain D of entities admitted as values of variables in L_T.

C1. *D* includes a denumerable subdomain *I* of entities.

C2. Any ordered n-tuple of entities in *D* (for any finite *n*) belongs also to *D*.

C3. Any class of entities in *D* belongs also to *D*.

I shall now indicate briefly how these conventions yield all the customary kinds of entities referred to in scientific theories. To facilitate the understanding, I shall first use the customary way of speaking and the customary terms for certain kinds of entities, and only later add a warning against a possible misinterpretation of these formulations.

First about mathematical entities. Since the subdomain *I* stipulated in *C1* is denumerable, we may regard its elements as the natural numbers 0, 1, 2, etc. If *R* is any relation whose members belong to *D*, then *R* may be construed as a class of ordered pairs of its members. Therefore, according to *C2* and *C3*, *R* belongs also to *D*. Now the (positive and negative) integers can, in the usual way, be constructed as relations of natural numbers. Thus, they belong also to *D*. Analogously, we proceed to rational numbers as relations among integers, to real numbers as classes of rational numbers, and to complex numbers as ordered pairs of real numbers. Furthermore, we obtain classes of numbers of these kinds, relations among them, functions (as special kinds of relations) whose arguments and values are numbers, then classes of functions, functions of functions, etc. Thus *D* includes all those kinds of entities needed in the purely mathematical part of L_T.

Now we proceed to physics. We assume that L_T is based upon a particular space-time coordinate system; thus the space-time points are ordered quadruples of real numbers and hence, according to *C2*, belong to *D*. A space-time region is a class of space-time points. Any particular physical system of which a physicist may speak, e.g., a material body or a process of radiation, occupies a certain space-time region. When a physicist describes a physical system or a process occurring in it or a momentary state of it, he ascribes values of physical magnitudes (e.g., mass, electric charge, temperature, electromagnetic field intensity, energy, and the like) either to the space-time region as a whole or to its points. The values of a physical magnitude are either real numbers or n-tuples of such. Thus a physical magnitude is a function whose arguments are either space-time points or regions and whose values are either real numbers or n-tuples of such. Thus, on the basis of our conventions, the domain *D* contains space-time points and regions, physical magnitudes and their values, physical systems and their states. A physical system itself is nothing else than a space-time region characterized in terms of magnitudes. In a similar way, all other entities occurring in physical theories can be shown to belong to *D*.

Psychological concepts are properties, relations, or quantitative magnitudes ascribed to certain space-time regions (usually human organisms or classes of such). Therefore they belong to the same logical types as concepts of physics, irrespective of the question of their difference in meaning and way of definition. Note that the logical type of a psychological concept is also independent of its methodological nature, e.g., whether based on observation of behavior or on introspection; philosophers seem sometimes not to realize this. Thus the domain *D* includes also all entities referred to in psychology. The same holds for all social sciences.

We have considered some of the kinds of entities referred to in mathematics, physics, psychology, and the social sciences and have indicated that they belong to the domain D. However, I wish to emphasize here that this talk about the admission of this or that kind of entity as values of variables in L_T is only a way of speaking intended to make the use of L_T, and especially the use of quantified variables in L_T, more easily understandable. Therefore the explanations just given must not be understood as implying that those who accept and use a language of the kind here described are thereby committed to certain "ontological" doctrines in the traditional metaphysical sense. The usual ontological questions about the "*reality*" (in an alleged metaphysical sense) of numbers, classes, space-time points, bodies, minds, etc., are pseudo questions

without cognitive content. In contrast to this, there is a good sense of the word "real," *viz.*, that used in everyday language and in science. It may be useful for our present discussion to distinguish two kinds of the meaningful use of "real," *viz.*, the common sense use and the scientific use. Although in actual practice there is no sharp line between these two uses, we may, in view of our partition of the total language L into the two parts L_O and L_T, distinguish between the use of "real" in connection with L_O, and that in connection with L_T. We assume that L_O contains only one kind of variable, and that the values of these variables are possible observable events. In this context, the question of reality can be raised only with respect to possible events. The statement that a specified possible observable event, e.g., that of this valley having been a lake in earlier times, is real means the same as the statement that the sentence of L_O which describes this event is true, and therefore means just the same as this sentence itself: "This valley was a lake."

For a question of reality in connection with L_T, the situation is in certain respects more complicated. If the question concerns the reality of an event described in theoretical terms, the situation is not much different from the earlier one: to accept a statement of reality of this kind is the same as to accept the sentence of L_T describing the event. However, a question about the reality of something like electrons in general (in contradistinction to the question about the reality of a cloud of electrons moving here now in a specified way, which is a question of the former kind) or the electromagnetic field in general is of a different nature. A question of this kind is in itself rather ambiguous. But we can give it a good scientific meaning, e.g., if we agree to understand the acceptance of the reality, say, of the electromagnetic field in the classical sense as the acceptance of a language L_T and in it a term, say '*E*,' and a set of postulates *T* which includes the classical laws of the electromagnetic field (say, the Maxwell equations) as postulates for '*E*.' For an observer X to "accept" the postulates of *T*, means here not simply to take *T* as an uninterpreted calculus, but to use *T* together with specified rules of correspondence C for guiding his expectations by deriving predictions about future observable events from observed events with the help of *T* and C.

I said previously that the elements of the basic domain *I* may be regarded as natural numbers. But I warned that this remark and the others about real numbers, etc., should not be taken literally but merely as a didactic help by attaching familiar labels to certain kinds of entities or, to say it in a still more cautious way, to certain kinds of expressions in L_T. Let the expressions corresponding to the domain I be "O," "O′," "O″," etc. To say that "O" designates the number zero, "O′" the number one, etc., gives merely the psychological help of connecting these expressions for the reader with useful associations and images, but should not be regarded as specifying part of the interpretation of L_T. All the interpretation (in the strict sense of this term, i.e., observational interpretation) that can be given for L_T is given in the C-rules, and their function is essentially the interpretation of certain sentences containing descriptive terms, and thereby indirectly the interpretation of the descriptive terms of V_T. On the other hand, the essential service that the expressions "O" etc. give, consists in the fact that they represent a particular kind of structure (*viz.*, a sequence with an initial member but no terminal member). Thus the structure can be uniquely specified but the elements of the structure cannot. Not because we are ignorant of their nature; rather, there is no question of their nature. But then, since the sequence of natural numbers is the most elementary and familiar example of the sequential structure here in question, no harm is done in saying that those expressions designate entities and that these entities are the natural numbers, as long as we are not misled by these formulations into asking metaphysical pseudo questions.

In the earlier discussion of the observation language L_O (Section II), we considered certain restrictive requirements, like those of nominalism, finitism, etc., and found them acceptable. However, the situation with respect to the theoretical language is entirely different. For L_T we do not claim to have a complete interpretation, but only the indirect and partial interpretation given by the correspondence rules. Therefore, we should feel free to choose the logical structure of this language as it best fits our needs for the purpose for which the language is constructed.

Thus here in L_T there is no reason against the three conventions, although their acceptance violates the first five requirements mentioned in Section II. First, before the C-rules are given, L_T, with the postulates T and the rules of deduction, is an uninterpreted calculus. Therefore the earlier requirements cannot be applied to it. We are free in the construction of the calculus; there is no lack of clarity, provided the rules of the calculus are clearly given. Then the C-rules are added. All they do is, in effect, to permit the derivation of certain sentences of L_O from

certain sentences of L_T or vice versa. They serve indirectly for derivations of conclusions in L_O, e.g., predictions of observable events, from given premises in L_O, e.g., reports of results found by observation, or for the determination of the probability of a conclusion in L_O on the basis of given premises in L_O. Since both the premises and the conclusion belong to L_O, which fulfills the restricting requirements, there can be no objection against the use of the C-rules and of L_T, as far as the meaningfulness of the results of the derivation procedure is concerned.

V. The Correspondence Rules C

There is no independent interpretation for L_T. The system T is in itself an uninterpreted postulate system. The terms of V_T obtain only an indirect and incomplete interpretation by the fact that some of them are connected by the rules C with observational terms, and the remaining terms of V_T are connected with the first ones by the postulates of T. Thus it is clear that the rules C are essential; without them the terms of V_T would not have any observational significance. These rules must be such that they connect sentences of L_O with certain sentences of L_T, for instance, by making a derivation in the one or the other direction possible. The particular form chosen for the rules C is not essential. They might be formulated as rules of inference or as postulates. Since we assume that the logical structure of the language is sufficiently rich to contain all necessary connectives, we may assume that the rules C are formulated as postulates. Let C be the conjunction of these *correspondence postulates*. As an example, we may think of L_T as a language of theoretical physics, based on a space-time coordinate system. Among the rules C there will be some basic ones, concerning space-time designations. They may specify a method for finding the coordinates of any observationally specified location, e.g., the method used by navigators for determining the position (the spatial coordinates: longitude, latitude, and altitude) and time. In other words, these C-rules specify the relation R which holds between any observable location u and the coordinates x, y, z, t, where x, y, z are the spatial coordinates and t is the time coordinate of u. More exactly speaking, the relation R relates to an observable space-time region u, e.g., an observable event or thing, a class u′ of coordinate quadruples which may be specified by intervals around the coordinate values x, y, z, t.

On the basis of these C-rules for space-time designations, other C-rules are given for terms of V_T, e.g., for some simple physical magnitudes like mass, temperature, and the like. These rules are spatiotemporally general, i.e., they hold for any space-time location. They will usually connect only very special kinds of value-distributions of the theoretical magnitude in question with an observable event. For example, a rule might refer to two material bodies u and v (i.e., observable at locations u and v); they must be neither too small nor too large for an observer to see them and to take them in his hands. The rule may connect the theoretical term "mass" with the observable predicate "heavier than" as follows: "If u is heavier than v, the mass of u′ (i.e., the mass of the coordinate region u′ corresponding to u) is greater than the mass of v′." Another rule may connect the theoretical term "temperature" with the observable predicate "warmer than" in this way: "If u is warmer than v, then the temperature of u′ is higher than that of v′."

As these examples show, the C-rules effect a connection only between certain sentences of a very special kind in L_T and sentences in L_O. The earlier view, that for some terms of V_T there could be definitions in terms of V_O, called either 'correlative definitions' (Reichenbach) or 'operational definitions' (Bridgman), has been abandoned by most empiricists as an oversimplification (see Section X). The essential incompleteness of the interpretation of theoretical terms was pointed out in my *Foundations of Logic and Mathematics* (6) and is discussed in detail by Hempel in (15, §3) and (16, §7). Moreover, it cannot be required that there is a C-rule for every term of V_T. If we have C-rules for certain terms, and these terms are connected with other terms by the postulates of *T*, then these other terms thereby also acquire observational significance. This fact shows that the specification, not only of the rules C, but also of the postulates T, is essential for the problem of meaningfulness. The definition of meaningfulness must be relative to a theory *T*, because the same term may be meaningful with respect to one theory but meaningless with respect to another.

In order to have a more concrete picture, we may think of the terms of V_T as quantitative physical magnitudes, e.g., functions from space-time-points (or finite space-time-regions) to real numbers (or n-tuples of real numbers). The postulates *T* may be conceived of as representing the fundamental laws of physics, not other physical statements, however well established. Let us think of the postulates *T* and the rules C as being completely general with respect to space and time—that is

as not containing references to any particular position in space or in time.

In the above examples, the C-rules have the form of universal postulates. A more general form would be that of statistical laws involving the concept of statistical probability (which means roughly, relative frequency in the long run). A postulate of this kind might say, for example, that, if a region has a certain state specified in theoretical terms, then there is a probability of 0.8 that a certain observable event occurs (which means that, on the average, in 80 per cent of those cases this event occurs). Or it might, conversely, state the probability for the theoretical property, with respect to the observable event. Statistical correspondence rules have so far been studied very little. (The probability conception of the psi-functions in quantum mechanics might perhaps be regarded as an example of probabilistic C-rules, as some customary formulations by physicists would suggest. I think, however, that this conception constitutes a probability connection *within* L_T rather than between L_T and L_O. What physicists often call "observable magnitudes," e.g., mass, position, velocity, energy, frequency of waves, and the like, are not "observable" in the sense customary in philosophical discussions of methodology, and therefore belong to the theoretical concepts in our terminology.) For the sake of simplicity, in most of my discussions here I shall think of the C-rules as postulates of universal form.

VI. A Criterion of Significance for Theoretical Terms

My task is to explicate the concept of the empirical meaningfulness of theoretical terms. I shall use "empirical significance" or, for short, "*significance*" as a technical expression for the desired explication. In preparation for the task of explication, let me try to clarify the explicandum somewhat more, i.e., the concept of empirical meaningfulness in its presystematic sense. Let 'M' be a theoretical term of V_T; it may designate a physical magnitude *M*. What does it mean for 'M' to be *empirically meaningful*? Roughly speaking, it means that a certain assumption involving the magnitude *M* makes a difference for the prediction of an observable event. More specifically, there must be a certain sentence S_M about *M* such that we can infer with its help a sentence S_O in L_O. (The inference may be either deductive, as I shall take it to be in the following discussion, or, more generally, probabilistic.) It is, of course, not required that S_O is derivable from S_M alone.

It is clear that we may use in the deduction the postulates T and the rules C. If now S_M contains not only 'M' but also other terms of V_T, then the fact that S_O is deducible does not prove that 'M' is meaningful, because this fact may just be due to the occurrence of the other terms. Therefore I shall require that S_M contain 'M' as the only term of V_T. Now it may be that any assumption involving only the magnitude M is in itself too weak to lead to an observational consequence, and that we have to add a second assumption S_K containing other terms of V_T but not 'M'. Let K be the class of these other terms. For example, S_M may say that, at a certain space-time point, M has the value 5, and S_K may say that, at the same space-time point or in its surroundings, certain other magnitudes have specified values. If S_O can be deduced from the four premises S_M, S_K, T, and C, while it cannot be deduced from S_K, T, and C alone, then the sentence S_M makes a difference for the prediction of an observable event, and therefore has observational significance. Since 'M' is the only descriptive term in S_M, 'M' itself has observational significance. However, this result must be qualified by a proviso. Since we have used the second assumption S_K involving the terms of K, we have shown only that 'M' is meaningful provided that the terms of K are meaningful. For this reason the definition of the significance of 'M' must be made relative not only to T and C, but also to the class K. 'M' is shown by the indicated procedure to be significant provided the terms of K have been found by a previous examination to be significant. Therefore the terms of V_T must be examined in a serial order. The first terms of V_T must be such that they can be shown to be significant without presupposing the significance of other descriptive terms. This will be the case for certain terms of V_T which are directly connected by C-rules with L_O. Other terms of V_T can then be shown to be significant by using the proved significance of the first terms, and so on. The total V_T can be regarded as significant only if we can show for a certain sequence of its terms that each term is significant relative to the class of the terms preceding it in the sequence.

It is clear that the definition must be relative to T, because the question whether a certain term in L_T is significant cannot possibly be decided without taking into consideration the postulates by which it is introduced. Perhaps the objection might be raised that, if significance is dependent upon T, then any observation of a new fact may compel us to take as nonsignificant a term so far regarded as significant or vice

versa. However, it should be noted first that the theory T which is here presupposed in the examination of the significance of a term, contains only the postulates, that is, the fundamental laws of science, and not other scientifically asserted sentences, e.g., those describing single facts. Therefore the class of the terms of L_T admitted as significant is not changed whenever new facts are discovered. This class will generally be changed only when a radical revolution in the system of science is made, especially by the introduction of a new primitive theoretical term and the addition of postulates for that term. And note further that the criterion here proposed is such that, although the whole of the theory T is presupposed in the criterion, the question of significance is still raised for each term separately, not only for the vocabulary V_T as a whole.

On the basis of the preceding considerations, I shall now give definitions for the concept of significance of descriptive terms in the theoretical language. The definition *D1* will define the auxiliary concept of *relative significance*, i.e., the significance of 'M' relative to a class K of other terms. Then the concept of significance itself will be defined in *D2*. According to our previous considerations, the concept of significance must furthermore be relative to the theoretical language L_T, the observation language L_O, the set of postulates T, and the correspondence rules C. We presuppose that the specifications of the languages L_T and L_O contain also a specification of the classes of descriptive terms, that is, V_T and V_O, respectively.

D1. A term 'M' is *significant relative to the class K* of terms, with respect to L_T, L_O, T, and $C =_{Df}$ the terms of K belong to V_T, 'M' belongs to V_T but not to K, and there are three sentences, S_M and S_K in L_T and S_O in L_O, such that the following conditions are fulfilled:

(a) S_M contains 'M' as the only descriptive term.
(b) The descriptive terms in S_K belong to K.
(c) The conjunction $S_M.S_K.T.C$ is consistent (i.e., not logically false).
(d) S_O is logically implied by the conjunction $S_M.S_K.T.C$.
(e) S_O is not logically implied by $S_K.T.C$.

The condition (c) is only added to assure that the situation described in S_M and S_K is possible, i.e., not excluded by the postulates T and the C-rules; otherwise the condition (d) would be trivially fulfilled.

D2. A term 'M_n' is *significant* with respect to L_T, L_O, T and $C =_{Df}$ there is a sequence of terms 'M_1', . . ., 'M_n' of V_T, such that every term 'M_i' ($i = 1, \ldots, n$) is significant relative to the class of those terms which precede it in the sequence, with respect to L_T, L_O, T, and C.

The sequence of terms referred to in *D2* must obviously be such that the first term 'M_1' can be shown to be significant without the help of other terms of V_T. In this case 'M_1' satisfies *D1*; the class *K* is the null class; the sentence S_K contains no descriptive terms; it is logically true and can therefore be omitted. In the simplest case of this kind, 'M_1' occurs in a C-rule, like "mass" and "temperature" in our previous examples. Suppose that the first three terms of our sequence are of the kind described. Then for the fourth term, the sentence S_K may contain any one or all three of these terms. In this way we may proceed, step by step, to other terms, which may be more and more remote from direct observation.

(A slightly stronger criterion might be taken into consideration, obtained by the following modification of *D1*. In addition to the sentence S_M, another sentence S'_M is used, which contains likewise '*M*' as the only descriptive term. Then the analogue to condition (c) for S'_M is added, and furthermore the analogue to condition (d) with S'_M taking the place of S_M and the negation of S_O taking the place of S_O. Thus here the assumption S_M leads to an observable consequence, as in *D1*, but another assumption S'_M about *M*, incompatible with S_M, leads to another observable consequence. However, the simpler criterion stated in *D1* seems sufficient as a minimum requirement for significance.)

In the informal discussion at the beginning of this section, I have referred to the *deduction* of S_O from certain premises. Correspondingly, *D1*(d) requires that S_O is logically implied by the premises. However, this simple situation holds only if the C-postulates have universal form, as we mostly assume in our discussions. In the more general case that also statistical laws are admitted as C-postulates (see the remark at the end of Section V) and perhaps also as postulates of *T*, then the result is a probability connection between $S_M.S_K$ on the one hand, and S_O on the other. In this case, the conditions (d) and (e) in *D1* are to be replaced by the condition that the probability of S_O relative to $S_M.S_K$, presupposing *T* and *C*, is different from the probability of S_O relative to S_K alone.

VII. The Adequacy of the Criterion of Significance

The criterion here proposed is admittedly very weak. But this is a result of the development of empiricism in these last decades. The original formulations of the criterion were found to be too strong and too narrow. Therefore, step by step, more liberal formulations were

introduced. Hempel has given in his article (15) a clear survey of this development. One change was the replacement of the principle of verifiability by the weaker requirement of confirmability or testability, as formulated in my paper (5). At the time of that paper, I still believed that all scientific terms could be introduced as disposition terms on the basis of observation terms either by explicit definitions or by so-called reduction sentences, which constitute a kind of conditional definition (see Section X). Today I think, in agreement with most empiricists, that the connection between the observation terms and the terms of theoretical science is much more indirect and weak than it was conceived either in my earlier formulations or in those of operationism. Therefore a criterion of significance for L_T must likewise be very weak.

In discussions of the requirement of confirmability (or, in earlier times, verifiability) the question was sometimes raised whether the possibility of the event which constitutes the confirming evidence was to be understood as logical possibility or as causal possibility (i.e., compatibility with the laws of nature or the laws of a given theory). According to Schlick's conception (22, p. 153) the possibility should be understood in the widest sense, as logical possibility. His main argument was the uncertainty about possibility in an empirical sense. He pointed out that the observer does not know whether certain operations are empirically possible for him or not. For example, he does not know whether he is able to lift this table; he is quite certain that he cannot lift an automobile; but both events are still conceivable and should therefore be regarded as possible evidence. Schlick's point was that a question of significance should never be dependent upon contingent facts.

On the other hand, Reichenbach and I (5, p. 423) maintained the view that logical possibility is not sufficient, but that physical (or, more generally, causal) possibility is required. The question whether a given sentence of L_T is confirmable must be taken as relative to a theory *T*. In examining such a question, a proposed evidence or a proposed test procedure could certainly not be accepted if they were incompatible with *T*. For example, on the basis of modern physics, which takes the velocity of light as the maximum signal velocity, any proposed test or evidence involving a signal with a higher velocity could not be accepted as proof of significance. The definition *D1* is based on this conception. The conjunction $S_M.S_K.T.C$ is required to be consistent by condition (c). Since S_O is logically implied by this conjunction, $S_M.S_K.S_O$ is compatible with T and C and thus causally possible. However, it is to be noted that causal possibility as here understood is much weaker than the kind of empirical possibility which Schlick had seemed to have in mind. In Schlick's example, neither the lifting of the table nor that of the automobile is excluded by our criterion, because these events are not incompatible with the T (and C); T contains only the fundamental laws of science, while those events are merely excluded by our empirical knowledge of the observer's ability to lift things.

I shall now examine the question of the adequacy of our criterion in more specific terms. Let us consider the case that the vocabulary V_T consists of two parts, V_1 and V_2, such that the terms of V_1 are empirically meaningful, while those of V_2 are entirely devoid of any empirical meaning. To make this presupposition about V_1 and V_2 more specific, we assume the following:

(1) If S_1 and S_2 are any sentences of L such that all descriptive terms of S_1 belong to V_1 or to the observational vocabulary V_O and those of S_2 to V_2, then neither of the two sentences logically implies the other, unless the implying sentence is logically false or the implied sentence is logically true.

Now a proposed criterion for the significance of terms of V_T should be regarded as too narrow if it excluded a term of V_1, and as too broad if it admitted a term of V_2. It would be adequate only if it were neither too narrow nor too broad.

For example, we might think of V_1 as containing terms of physics, and of V_2 as containing meaningless terms of speculative metaphysics such that the supposition (1) holds.

First let us consider a postulate system T' consisting of two parts, T'_1 and T'_2, T'_1 containing only terms of V_1, and T'_2 only terms of V_2. T'_1 may, for example, consist of fundamental laws of physics, and T'_2 of metaphysical principles. A criterion of significance which is adequate in this special case can easily be given. We call a postulate of a system T an *isolated postulate* if its omission from T does not diminish the class of sentences in L_O which are deducible from T with the help of the C-rules. Then we take a term of V_T as significant if it occurs in a C-rule or in a non-isolated postulate of T. In the case of the above system T', according to (1), all postulates of T'_2 and no others are isolated; therefore all terms of V_1 and no others fulfill the criterion of significance just mentioned.

This criterion is, however, not generally adequate. It would, for example, not work for a theory T'' logically equivalent to T' but such that no postulate of T'' is isolated. Those who are sceptical about the possibility of a criterion of significance for L_T have probably a situation of this kind in mind. (Hempel discusses a similar example.) They believe that it is not possible to give a criterion for postulate systems like T''. However, I think that the criterion for terms proposed in Section VI is adequate for cases of this kind. Consider for the postulate system T'' the sequence of terms which is required in *D2*. This sequence must necessarily begin with physical terms of V_1, because, according to our assumption (1), there are no C-rules for any of the metaphysical terms of V_2. Then the sequence may go on to further physical terms, which are connected with L_O not directly by C-rules, but indirectly by other physical terms. Now we shall see that the sequence cannot reach any term of V_2; thus our criterion is not too broad for systems like T''. We will show this by an indirect proof. We assume that the sequence reaches terms of V_2; let 'M' be the first term of V_2 in the sequence; hence the preceding terms belong to V_1, and thus are meaningful. 'M' is significant relative to the class K of the preceding terms, with respect to L_T, L_O, T'', and C, in the sense of *D1*. Intuitively speaking, 'M' must then be meaningful, in contradiction to our presupposition about V_2. Our task is, to derive formally a contradiction with the presupposition (1).

According to *D1*(d):

(2) $S_M.S_K.T''.C \supset S_O$ is logically true.

Now T'' is logically equivalent to T' and thus to $T'_1.T'_2$. Hence we obtain from (2) with a simple transformation:

(3) $S_M.T'_2 \supset U$ is logically true, where U is $S_K.T'_1.C \supset S_O$.

Hence:

(4) $S_M.T'_2$ logically implies U.

Now all descriptive terms in $S_M.T'_2$ belong to V_2, and those in U belong to V_1 or V_O. Thus (4) is in contradiction to (1), because

(5) $S_M.T'_2$ is not logically false (by *D1*(c)), and

(6) U is not logically true (by *D1*(e)).

This shows that the sequence cannot reach the terms of V_2.

We have shown that our criterion is not too broad if the given set of postulates T'' is logically equivalent to a set T' which consists of two parts, one containing only meaningful terms of V_1, the other only meaningless terms of V_2. The situation would be different for a theory T that did not fulfill this condition. In this case, T must include a postulate A such that A contains terms from both V_1 and V_2, but A is not logically equivalent to a conjunction $A_1.A_2$ in which A_1 contains only terms of V_1, and A_2 only terms of V_2. But such a postulate A would express a genuine connection between the occurring terms of V_2 and those of V_1. Therefore these terms of V_2 would not be entirely devoid of empirical meaning, against our assumption.

The result that our criterion of significance is not too broad depends essentially on the following feature of our definitions. We refer in *D2* to a sequence of terms, and we require in effect for the significance of a term '*M*' of the sequence that '*M*' is significant (in the sense of *D1*) relative to the class *K* of the terms which precede '*M*' in the sequence and which therefore have already been found to be significant. We can easily see that the criterion would become too broad if we were to change *D2* so as to give up the requirement just mentioned. More specifically, we can show the following. A meaningless term 'M_2' of V_2 can, according to *D1*, be significant relative to a class *K* which contains, in addition to terms of V_1, also a meaningless term of V_2 different from 'M_2', say 'M'_2.' We shall show this first informally. The decisive point is that now, in distinction to our actual definition *D2*, we can have as the additional assumption S_K a sentence connecting the meaningless term 'M'_2' with a meaningful (physical) term of V_1, say 'M_1.' Now there may be a (metaphysical) postulate A_2 of T which connects M_2 with M'_2. With the help of this postulate, we can derive from the assumption S_M about M_2 alone a sentence about M'_2; from this with the sentence S_K mentioned above a physical sentence about M_1, and from this with a suitable C-rule an observation sentence.

The formal derivation is as follows. We take as a postulate of T:

(A_2) For every space-time point, the value of M'_2 is higher than that of M_2 by one.

We take as an instance of a C-rule:

$$(C_1)\ M_1(a') = 5 \supset S_O,$$

where a′ is the set of coordinates corresponding to the location a referred to in S_O. Finally we take S_K and S_M as follows:

$$(S_K)\ M_1(a') = M'_2(a'),$$

$$(S_M)\ M_2(a') = 4.$$

Now we can derive from S_M with A_2:

(i) $M'_2(a') = 5,$

hence with S_K:

(ii) $M_1 a' = 5,$

and hence with C_1:

(iii) S_O.

Thus the condition (d) in *D1* is fulfilled. Therefore, 'M_2' is significant relative to the class *K* of the terms 'M_1' and 'M'_2'.

We have just seen that, in the definition of the significance of '*M*' relative to *K*, we must not admit a meaningless term in *K* and thereby in the additional assumption S_K, because otherwise an observation sentence could be derived, leading to a deceptive appearance of significance. This is indeed excluded by *D2*. However, *D1* allows other premises for the derivation which contain meaningless terms, *viz.*, postulates of *T*. Not only the postulates which contain the meaningful terms of V_1 and the term '*M*' in question are allowed but also postulates containing any terms of V_2. Could this not lead to the same false appearance of significance for an actually meaningless term '*M*' as the use of meaningless terms in S_K would do? In the above example, S_K connected a meaningless term 'M'_2' with a meaningful term 'M_1', and this fact led to the undesired result. Now the use of *T* would lead to the same result if a postulate of *T* were to make a connection between those terms. For example, a postulate might yield as an instance the sentence "$M_1(a') = M'_2(a')$" which was used as S_K in the earlier example. Thus the same observation sentence S_O could be derived from S_M even without the use of any second assumption S_K. As an alternative, a postulate might state a connection between 'M'_2' and 'M_1' in a conditional form, which, though weaker, would likewise make possible a derivation of an observation sentence. Does then the fact that *D1* permits the use of all postulates *T* make this definition inadequate? It does not, because the occurrence of a postulate making a genuine connection between a term of V_1 and one of V_2 is excluded by our presupposition that the terms of V_1 are meaningful and those of V_2 meaningless. By virtue of such a postulate, the term of V_2 (in the example, 'M'_2') would obtain some measure of empirical meaning, as we observed earlier in this section with reference to the postulate A. The essential difference between the two cases is the following. If a sentence connecting a meaningful term with another term in an inseparable way (e.g., by an equation, a conditional, a disjunction or the like, in distinction to a conjunction, which can be separated into its components) is a postulate or provable on the basis of postulates, then it is stated as holding with physical necessity; therefore it conveys some empirical meaning on the second term. On the other hand, if the same sentence is not provable but is merely used as the additional assumption S_K in *D1*, then it has no such effect; it need not even be true.

The preceding considerations have shown that our criterion of significance, formulated in *D1* and *D2*, is not too liberal. It does not admit a term entirely devoid of empirical meaning. Now we shall consider the question whether the criterion might be too narrow. Suppose that the term '*M*' has some empirical meaning. Then it will be possible to derive an observation sentence from a suitable assumption *S* involving '*M*' and other terms. Could it then still happen that our criterion would exclude '*M*'? The definitions *D1* and *D2*, while permitting the inclusion of all postulates *T* and *C* among the premises for the derivation of the observation sentence, allow in addition only the two sentences S_K and S_M, for which specific restrictions are stated, especially the following:

(1) S_K may contain only terms of V_T which are different from '*M*' and have to be significant; hence the following terms are not allowed in S_K:

(a) terms of V_2,
(b) terms of V_O,
(c) The term '*M*'.

(2) S_M contains '*M*' as the only descriptive term.

We will now examine whether these restrictions are narrower than is necessary and thus might lead to the exclusion of a meaningful term '*M*.'

1a. We found earlier that it is necessary to exclude the terms of V_2 from S_K, because otherwise the criterion would become too broad.

1b. Is it necessary to exclude the observational terms V_O from the premises? Could it not be that, for the derivation of an observational conclusion S_O from S_M, we need, in addition to *T* and *C* and the assumption S_K in theoretical terms, some assumption in observation terms, say S'_O? This might well happen. But then the conditional sentence $S'_O \supset S_O$ is derivable from the premises specified in *D1*, and this is a sentence in L_O. Thus '*M*' would fulfill *D1*, with the conditional sentence taking the place of S_O.

1c and *2*. The condition (a) in *D1* requires that S_M contain '*M*' as the only descriptive term. The question might be raised whether this requirement is not too strong. Could not the following situation occur? '*M*' and the terms of *K* are meaningful, and S_O can indeed be derived with the help of *T* and *C* from an assumption *S* containing no other descriptive terms than '*M*' and the terms of *K*, but *S* cannot be split up into two sentences S_M and S_K such that S_M contains only '*M*' and S_K does not contain '*M*.' Let us assume that the sentence *S* refers to space-time points of a certain spatiotemporal region *a'*. Then we can form sentences S_M and S_K which fulfill the requirements of *D1* in the following way. Since *S* is supposed to be compatible with *T* and *C*, there must be a possible distribution of values of *M* for the space-time points of the region a', which is compatible with *T*, *C*, and *S*. Let '*F*' be a logical constant, designating a mathematical function which represents such a value distribution. Then we take the following sentence as S_M: "For every space-time point in a', the value of *M* is equal to that of *F*." This sentence S_M is compatible with *T.C.S*. Then we take as S_K the sentence formed from *S* by replacing the descriptive term '*M*' by the logical constant '*F*'. Then S_M contains '*M*' as the only descriptive term and S_K contains only terms of *K*. Furthermore, *S* is logically implied by S_M, and S_K. S_O is logically implied by *S.T.C.*, according to our assumption, and hence also by $S_M.S_K.T.C$. Therefore '*M*' fulfills the definition *D1*.

Thus we have not found a point in which our criterion is too narrow.

VIII. A Criterion of Significance for Theoretical Sentences

The following two problems are closely connected with each other: first, the problem of a criterion of significance for descriptive constants, and second, the problem of the logical forms to be admitted for sentences. For the theoretical language, the connection between these problems is still closer than for the observation language. In the latter, we may decide to have primitive predicates like "blue," "cold," "warmer than," and the like, while we are still undecided as to the forms of sentences, especially of general sentences, and the structure of the logic to be built into the language. On the other hand, if we wish to have terms like "temperature," "electromagnetic field," etc. as primitives in L_T, then we need also the accepted postulates for them, and thus we have to admit real number expressions, general sentences with real number variables, etc.

It seems to me that the best approach to the problem of a criterion of significance for sentences is the following. We look first for solutions to the two problems mentioned above; and then we take the most liberal criterion of significance for sentences which is compatible with those solutions. That is to say, we then accept as a significant sentence any expression that has any of the admitted logical forms and contains only descriptive constants which are significant. (I have used a similar approach for L_O in (5).) I propose to apply this procedure now to L_T.

A criterion of significance for descriptive terms was given in Section VI. Some of the questions concerning the logical forms of sentences were discussed in Section IV, especially the question of the kinds of variables to be admitted in universal and existential quantifiers. We decided to admit at least those kinds of variables and forms of sentences which are essential for classical mathematics. Without actually specifying here the details of the rules, we shall now assume that the logical forms of sentences have been chosen on the basis of the considerations in Section IV, and that the rules of formation for L_T have been laid down in accordance with this choice. Then, applying the procedure proposed above, we define as follows:

D3. An expression A of L_T is a *significant sentence* of $L_T =_{Df}$

(a) A satisfies the rules of formation of L_T,

(b) every descriptive constant in A is a significant term (in the sense of D2).

The procedure used in this definition might perhaps appear as obvious. However, a closer examination shows that this is not the case. In fact, this form of the definition (aside from the question of its content, i.e., the choice of the particular rules of formation and of the particular significance criterion for terms) is not in agreement with certain very narrow criteria of significance which were sometimes proposed. For

example, verifiability as a condition for the significance of a sentence was sometimes understood in the strict sense of the actual possibility of carrying out a procedure which would lead either to a verification or a falsification of the sentence. According to this criterion, in contrast to D3, the significance of a sentence is not only dependent upon its logical form and the nature of the descriptive constants occurring in it, but also upon the location in space and time referred to and the development of technology. For example, an empiricist applying this narrow criterion would regard as significant a sentence ascribing an observable property *P* to a body in his laboratory, while he would reject as nonsignificant another sentence which ascribes the same property to a body not accessible to him or not accessible to any human being, e.g., because of technical difficulties or remoteness in space or time.

However, even at the time of the Vienna Circle, we did not interpret the principle of verifiability in this narrow sense. We emphasized that the principle required, not the actual possibility of determination as true or false, but only the possibility *in principle*. By this qualification we intended to admit cases in which the determination was prevented only by technical limitations or by remoteness in space or time. We accepted, for example, a sentence about a mountain on the other side of the moon as meaningful. We stated the general rule that, if a description of an event in our neighborhood is regarded as meaningful, then an analogous description of an event in prehistoric times, or an event on the earth before there were human beings, or before there were any organisms, or at a future time when human beings will not exist any more, should likewise be accepted as meaningful. On the basis of this conception, the space-time location referred to in a sentence was regarded as irrelevant for the question of meaningfulness; this is in accord with D3.

If D3 is accepted and, in line with our earlier considerations in Section IV, all constants, variables and forms of sentences of classical mathematics are admitted in L_T, then the class of significant sentences of L_T is very comprehensive. We must realize that it includes certain sentences for which no observational evidence can ever be relevant, e.g., the sentence: "The value of the magnitude *M* at a certain space-time point is a rational number," where '*M*' is significant. But every physicist would reject a language of physics so restricted that sentences of this and similar kinds were excluded. He would regard their inclusion as a negligible price to be paid for the great convenience of using the whole of classical mathematics. It seems to me that no serious objections can be raised against these sentences, since it is in any case not possible to give an observational interpretation for more than a small part of the sentences of L_T. We should require no more than that for such a magnitude there are certain sentences which have an influence on the prediction of observable events and thus the magnitude itself has some amount of observational meaning.

I wish to emphasize that the proposed criterion for the significance of sentences is not meant to guarantee the fruitfulness of *T*. If all terms of V_T fulfill *D2* and the postulates *T* are in accord with the rules of formation, then these postulates are indeed regarded as significant. But this should by no means be understood as implying that *T* must then be a scientifically satisfactory theory. *T* may still contain postulates which are of very little use from a scientific point of view. But the question of scientific fruitfulness of sentences and of a theory should be clearly distinguished from the question of empirical significance. There is no sharp boundary line between fruitful and useless hypotheses or theories: this is rather a matter of degree. It seems even doubtful whether it is possible to formulate in a completely general way a definition of a quantitative degree of fruitfulness of a scientific theory.

It should be noted that the significance criterion for L_T cannot be simply absorbed into the rules of formation. These rules determine only the forms of sentences, not the choice of primitive descriptive terms. The significance of these terms depends on other rules of L_T, viz., the list of postulates *T* and of C-postulates and the rules of logical deduction, as a glance at the essential condition (d) in *D1* shows. (The rules of deduction may be given either in a syntactical form, as rules of derivation in a calculus, or in a semantical form, in terms of logical implication. I have used in *D1* the latter form because it is more comprehensive; it presupposes rules specifying models and ranges, not given in this article.)

IX. Disposition Concepts

Among the descriptive terms which do not belong to the observation language L_O there are two different kinds, which today, in distinction to my previous conception, I should like to regard as essentially different. One kind is that of the theoretical terms, which we have

discussed in detail in this article. The other kind I will call (pure) disposition terms. In my view, they occupy an intermediate position between the observational terms of L_O and the theoretical terms; they are more closely related to the former than to the latter. The name 'observation language' may be understood in a narrower or in a wider sense; the observation language in the wider sense includes the disposition terms. In this article I take the observation language L_O in the narrower sense. All primitive predicates in this language designate directly observable properties or relations of observable things or events; and a nonprimitive term is admitted in L_O only if it can be defined on the basis of the primitive terms by an explicit definition in an extensional form, that is, not involving either logical or causal modalities. The *extended observation language* L'_O is constructed from the original observation language L_O by the addition of new terms in a way now to be described. Suppose that there is a general regularity in the behavior of a given thing of such a kind that, whenever the condition *S* holds for the thing or its environment, the event *R* occurs at the thing. In this case we shall say that the thing has the disposition to react to *S* by *R*, or for short, that it has the property D_{SR}. For example, elasticity is a disposition of this kind; a thing is called elastic if it shows the following regularity: whenever it is slightly deformed and then released (*S*), it resumes its original form (*R*). Or, an animal may have the disposition to react to a light in an otherwise dark environment (*S*), by approaching the light (*R*). Thus, *S* is sometimes a stimulus, and *R* is the response characteristic for the disposition in question (if we allow ourselves to use the terms 'stimulus' and 'response' not only in their literal sense applied to certain processes in organisms, as in the last example, but in a wider sense also to processes with inorganic bodies). When both *S* and *R* are specified, then the disposition concept D_{SR} is thereby completely characterized in its meaning. If both *S* and *R* can be described in L'_O, then we admit the introduction of the disposition term 'D_{SR}' as a new predicate in L'_O. The introduction of the first disposition terms in L'_O must be of such a kind that in each case both *S* and *R* are expressible in L_O. But once some disposition terms have been introduced in this way, then further disposition terms may be introduced in such a way that *S* and *R* are described by using not only the terms of L_O, but also the previously introduced disposition terms of L'_O.

(We will not discuss here the possible forms for the rule by which a disposition term is introduced on the basis of given S and R. This involves some technicalities which are not necessary for our present discussions. I will only mention two different forms for such rules that have been proposed. The first consists of so-called reduction sentences, which I proposed in (5). They represent a kind of conditional definition which uses only truth-functional connectives, but no modalities. The other method uses an explicit definition of a special form, involving logical and causal modalities; the exact form of definitions of this kind is at present not yet sufficiently clarified, but still under discussion.)

Sometimes multiple dispositions are used: $D_{S_1R_1,S_2R_2,\ldots,S_nR_n}$ is the disposition to react to S_1 by R_1, to S_2 by R_2, . . ., and finally to S_n by R_n. (In (5) I proposed to introduce a concept of this kind by several pairs of reduction sentences.) However, it seems preferable to admit only simple dispositions. Something similar to a multiple disposition can still be expressed by a conjunction of simple dispositions. Bridgman has emphasized that, strictly speaking, for one concept not more than one test procedure must be given. If we specify, say for "electric charge," three test procedures, then thereby we have given operational definitions for three different concepts; they should be designated by three different terms, which are not logically equivalent. As far as disposition concepts are concerned, in distinction to theoretical terms, I would agree with Bridgman in this point.

Let us now consider an important special kind of disposition. Let L''_O be that sublanguage of L'_O, in which the introduction of a disposition term 'D_{SR}' is permitted only if S and R are such that the observer is able to produce the condition S at will (at least in suitable cases), and that he is able to find out by suitable experiments whether the event R does or does not occur. In this case, by specifying S and R, a *test procedure* for the disposition D_{SR} is given. This procedure consists in producing the *test condition* S and then finding out whether or not the *positive test result* R occurs. If the observer finds for a given thing a sufficient number of positive instances, in which S is followed by R, and no negative instances, i.e., S followed by non-R, he may inductively infer that the general regularity holds and thus that the thing possesses the disposition D_{SR}. Let us call a disposition of this kind a "testable disposition." The class of *testable properties* includes observable properties and testable dispositions. All predicates in L''_O designate testable properties. The manipulations by which the experimenter produces the

test condition S are sometimes called *test operations*. The introduction of D_{SR} by a specification of the test operations and the characteristic result R is therefore sometimes called an *operational definition*. There is actually no sharp line between observable properties and testable dispositions. An observable property may be regarded as a simple special case of a testable disposition; for example, the operation for finding out whether a thing is blue or hissing or cold, consists simply in looking or listening or touching the thing, respectively. Nevertheless, in the reconstruction of the language it seems convenient to take some properties, for which the test procedure is extremely simple (as in the three examples just mentioned), as directly observable and use them as primitives in L_O.

The view has often been maintained, especially by empiricists, that only terms of the kind just described, may be regarded as empirically meaningful. Thus testability was taken as a criterion of significance. The *principle of operationism* says that a term is empirically meaningful only if an operational definition can be given for it. The requirements of testability and of operationism as represented by various authors are closely related to each other, differing only in minor details and in emphasis. (In my simplifying account above they even appear as identical.) The principle of operationism, which was first proposed in physics by Bridgman and then applied also in other fields of science, including psychology, had on the whole a healthy effect on the procedures of concept formation used by scientists. The principle has contributed to the clarification of many concepts and has helped to eliminate unclear or even unscientific concepts. On the other hand, we must realize today that the principle is too narrow.

That the requirements of testability and of operationism exclude some empirically meaningful terms, can easily be seen. Suppose that 'S' and 'R' are both testable and hence accepted as meaningful by a scientist who takes testability as a criterion of significance. Since now the meaning of the term 'D_{SR}' is given by the specification of S and R, there cannot be any good reason for him to reject this term as meaningless, even if the condition S cannot be produced at will. In the latter case, D_{SR} is not testable; but S may still occur spontaneously and then, by finding R or non-R, the observer may determine whether or not D_{SR} holds. Thus it seems preferable not to impose the restriction as in L''_O, but to allow the general procedure as in L'_O: we start with observable properties and allow the introduction of any disposition D_{SR}, provided that S and R are already expressible in our language L'_O.

(In (5), I gave an example of a meaningful but not testable term (p. 462) of the kind just described. I expressed there (§27) my preference for the more general procedure (as in L'_O) in comparison with that restricted by the requirement of testability (as in L''_O). Later it became clear by the consideration of theoretical concepts (see the next section of this paper) that a far more extensive liberalization of operationism is needed; this was emphasized by Feigl in (7) and (10) and by Hempel in (16) and (17).)

X. The Difference between Theoretical Terms and Pure Disposition Terms

I think today that, for most of the terms in the theoretical part of science and especially in physics, it is more adequate and also more in line with the actual usage of scientists, to reconstruct them as theoretical terms in L_T rather than as disposition terms in L'_O. The choice of the form of reconstruction depends to some extent upon the interpretation which we wish to give to the term, and this interpretation is not uniquely determined by the accepted formulations in science. The same term, say "temperature," may be interpreted, as I do interpret it, in such a way that it cannot be represented in L'_O but only in L_T; and, on the other hand, it may also be interpreted, e.g., by an operationist, in such a way that it fulfills the requirement of operationism. I shall now explain the reasons for my present view, which differs from that stated in (5).

A disposition term like 'D_{SR}' introduced by the general method described in the last section (for L'_O) may be called a "pure disposition term" in order to emphasize that it has the following characteristic features which distinguish it from terms in L_T:

1. The term can be reached from predicates for observable properties by one or more steps of the procedure described.

2. The specified relation between S and R constitutes the whole meaning of the term.

3. The regularity involving S and R, on which the term is based, is meant as universal, i.e., holding without exception.

The first characteristic distinguishes a pure disposition term like 'D_{SR}' from other disposition terms which are analogous to 'D_{SR}' but such that

the condition S and the characteristic result R are formulated in L_T rather than in L_O or L'_O. (They might be called "theoretical disposition terms"; we shall not discuss them further.) The second characteristic distinguishes 'D_{SR}' from any theoretical term because the latter is never completely interpreted. In (5) I recognized this "open" character of scientific terms, that is, the incompleteness of their interpretation. At that time I tried to do justice to this openness by admitting the addition of further dispositional rules (in the form of reduction sentences; see my remarks in Section IX above on multiple dispositions). I think now that the openness is more adequately represented in L_T; whenever additional C-rules or additional postulates are given, the interpretation of the term may be strengthened without ever being completed.

The third characteristic leads to the following important consequence: (i) If the thing *b* has the disposition D_{SR} and the condition S is fulfilled for *b*, then it follows logically that the result R holds for *b*.

Therefore:

(ii) If S holds for *b*, but R does not, then *b* cannot have the disposition D_{SR}. Thus, from a premise in L'_O not involving D_{SR}, at least a negative sentence about D_{SR} is derivable. For a theoretical term, say '*M*,' the situation is different. Let S_M be a sentence containing '*M*' as the only descriptive term. In the situation described in *D1* in Section VI, S_O is derivable from S_M and S_K (with the help of *T* and *C*, which may be regarded as belonging to the rules of L_T), and therefore non-S_M is derivable from non-S_O and S_K. Since S_K is not translatable into L_O or L'_O, the situation is here different from that in (ii). It is true that, for a term '*M*' occurring in a C-rule, there are sentences S_M and S_O such that S_O is derivable from S_M alone without the need of a second premise S_K; and hence non-S_M is derivable from non-S_O, so that the situation is similar to that in (ii). However, this holds only for sentences of a very special kind. Most of the sentences about *M* alone, even if '*M*' is a term occurring in a C-rule, are such that no C-rule is directly applicable, and therefore the derivation of an observation sentence is more indirect and needs additional premises in L_T, like S_K. Consider, for example, the term "mass," which is one of the physical terms most closely related to observational terms. There may be C-rules for "mass" (see the example in Section V). But no C-rule is directly applicable to a sentence S_M ascribing a certain value of mass to a given body, if the value is either so small that the body is not directly observable or so large that the observer cannot manipulate the body. (I mentioned in Section V the possibility of probabilistic C-rules. If all C-rules have this form, then no theoretical sentence is deducible from sentences in L_O or L'_O. Thus in a language of this kind, the difference between pure disposition terms and theoretical terms becomes still more striking.)

We have seen that pure disposition terms and theoretical terms are quite different in their logical and methodological characteristics. To which of these two kinds do scientific terms belong? For the terms of theoretical physics, both conceptions are represented among leading physicists. Bridgman interprets them in such a manner that they fulfill the requirement of operationism and thus are essentially pure dispositions. On the other hand, Henry Margenau emphasizes the importance of the method of introducing these terms by postulates and connecting only certain statements involving them with statements about observables; in this conception they are theoretical terms.

It seems to me that the interpretation of scientific terms as pure dispositions cannot easily be reconciled with certain customary ways of using them. According to (ii), the negative result of a test for a disposition must be taken as conclusive proof that the disposition is not present. But a scientist, when confronted with the negative result of a test for a certain concept, will often still maintain that it holds, provided he has sufficient positive evidence to outbalance the one negative result. For example, let I_o be the property of a wire carrying at the time t_o no electric current of more than 0.1 ampere. There are many test procedures for this property, among them one in which the test condition S consists in bringing a magnetic needle near to the wire, and the characteristic result R is the fact that the needle is not deflected from its normal direction by more than a certain amount. Suppose that the observer assumes from the arrangement of the experiment that I_o holds, e.g., because he does not see any of the ordinary sources of a current and he has obtained, in addition, positive results by some other tests for I_o (or for a physically equivalent property). Then it may be that he does not give up the assumption of I_o even if the above mentioned test with S and R leads to a negative result, that is, a strong deflection of the needle. He may maintain I_o because it is possible that the negative result is due to an unnoticed disturbing factor; e.g., the deflection of the needle may be caused by a hidden magnet rather than by a current in the wire. The fact that the scientist still assumes I_o in

spite of the negative result, *viz.*, S and non-R, shows that he does not take I_o as the pure disposition D_{SR} characterized by S and R, because, according to (ii), this disposition is logically incompatible with the negative result. The scientist will point out that the test procedure for I_o based on S and R should not be taken as absolutely reliable, but only with the tacit understanding "unless there are disturbing factors" or "provided the environment is in a normal state." Generally, the explicit or implicit inclusion of such an escape clause in the description of a test procedure for a concept M in terms of a condition S and a result R shows that M is not the pure disposition D_{SR}. Also, the name "operational definition" for the description of the test procedure is in this case misleading; a rule for the application of a term that permits possible exceptions should not be called a "definition" because it is obviously not a complete specification of the meaning of the term.

On the other hand, if the term in question, e.g., 'I_o', is a theoretical term, then the description of the test procedure involving S and R may well admit of exceptions in case of unusual disturbing factors. For example, it may be possible to derive from the postulates T, the C-rules, and factual premises about usual circumstances in a laboratory the conclusion that, if there is no strong current, there will not be a strong deflection of the needle, except in the case of unusual circumstances like a magnetic field from another source, a strong current of air, or the like.

Thus, if a scientist has decided to use a certain term 'M' in such a way, that for certain sentences about M, any possible observational results can never be absolutely conclusive evidence but at best evidence yielding a high probability, then the appropriate place for 'M' in a dual-language system like our system L_O-L_T is in L_T rather than in L_O or L'_O.

XI. Psychological Concepts

The method of reconstructing the language of science by the dual schema consisting of the observation language L_O and the theoretical language L_T and the distinction between pure dispositions and theoretical concepts were so far in this article illustrated mostly by examples taken from physics. In the historical development of science, physics was indeed the field in which the method of introducing terms by postulates without a complete interpretation was first used systematically. The beginning phase of this development may perhaps be seen in the classical mechanics of the eighteenth century; its character became more clearly recognizable in the nineteenth century, especially in the Faraday-Maxwell theory of the electromagnetic field and the kinetic theory of gases. The greatest and most fruitful application is found in the theory of relativity and in quantum theory.

We see at present the beginnings of similar developments in other fields of science, and there can be no doubt that here too the more comprehensive use of this method will lead in time to theories much more powerful for explanation and prediction than those theories which keep close to observables. Also in psychology, in these last decades, more and more concepts were used which show the essential features of theoretical concepts. The germs of this development can sometimes be found in much earlier periods and even, it seems to me, in some prescientific concepts of everyday language, both in the physical and psychological field.

In psychology still more than in physics, the warnings by empiricists and operationists against certain concepts, for which no sufficiently clear rules of use were given, were necessary and useful. On the other hand, perhaps due to the too narrow limitations of the earlier principles of empiricism and operationism, some psychologists became overcautious in the formation of new concepts. Others, whose methodological superego was fortunately not strong enough to restrain them, dared to transgress the accepted limits but felt uneasy about it. Some of my psychologist friends think that we empiricists are responsible for the too narrow restrictions applied by psychologists. Perhaps they overestimate the influence that philosophers have on scientists in general; but maybe we should plead guilty to some extent. All the more should we now emphasize the changed conception which gives much more freedom to the working scientist in the choice of his conceptual tools.

In a way similar to the philosophical tendencies of empiricism and operationism, the psychological movement of Behaviorism had, on the one hand, a very healthful influence because of its emphasis on the observation of behavior as an intersubjective and reliable basis for psychological investigations, while, on the other hand, it imposed too narrow restrictions. First, its total rejection of introspection was unwarranted. Although many of the alleged results of introspection were indeed questionable, a person's awareness of his own state of imagining, feel-

ing, etc., must be recognized as a kind of observation, in principle not different from external observation, and therefore as a legitimate source of knowledge, though limited by its subjective character. Secondly, Behaviorism in combination with the philosophical tendencies mentioned led often to the requirement that all psychological concepts must be defined in terms of behavior or behavior dispositions. A psychological concept ascribed to a person X by the investigator Y either as a momentary state or process or as a continuing trait or ability, was thus interpreted as a pure disposition D_{SR} of such a kind that S was a process affecting a sensory organ of X but observable also by Y, and R was a specified kind of behavior, likewise observable by Y. In contrast to this, the interpretation of a psychological concept as a theoretical concept, although it may accept the same behavioristic test procedure based on S and R, does not identify the concept (the state or trait) with the pure disposition D_{SR}. The decisive difference is this: on the basis of the theoretical interpretation, the result of this or of any other test or, generally, of any observations, external or internal, is not regarded as absolutely conclusive evidence for the state in question; it is accepted only as probabilistic evidence, hence at best as a reliable indicator, i.e., one yielding a high probability for the state.

In analogy to what I said in the previous section about physical terms, I wish to emphasize here for psychological terms that their interpretation as pure disposition terms is not in itself objectionable. The question is only whether this interpretation is in accord with the way the psychologist intends to use the term, and whether it is the most useful for the purpose of the whole of psychological theory, which is presumably the explanation and prediction of human behavior. Suppose that the psychologist Y declares that he understands the term "an IQ higher than 130" in the sense of the pure disposition D_{SR} to react to a specified kind of test S by a response of a specified kind R, where S and R are specified in terms of overt behavior. He is free to choose this interpretation provided he is consistent in it and willing to accept its implications. Suppose that he assumes on the basis of ample previous evidence that (at present) the person X has an IQ higher than 130. Then, due to his interpretation, he is compelled to give up the assumption if today the test result is negative, i.e., X's response to the test S is not of the specified kind R. (This follows from (ii) in Section X.) He cannot even re-accept the assumption later when he learns that during the test X was in a very depressed mood, which, however, he neither admitted on question nor showed in his behavior at the time of the test. Can the psychologist not escape from this embarrassing consequence by saying that X's later admission of his depressed state showed that the condition S was actually not fulfilled? Not easily. There would have to be a rule as part of the specification of S that would enable him to make the exception. Let us consider three possibilities for a rule.

1. Let the rule merely say that, at the time t_o of the test, there must be first a complete lack of any observable sign of a disturbed emotional state at time t_o and second a negative answer to a question about such a state. Here the condition S was actually fulfilled and thus the psychologist has no way out.

2. Let the rule add, moreover, that also at no later time must there be a sign indicating a disturbance at time t_o. In this case, S was indeed not fulfilled. But a test procedure containing a rule of this kind would be practically useless, because it could never be completed before the death of the person.

3. Finally, let the rule refer not to behavioral signs but to the emotional state itself. Here the test procedure is not a strictly behavioristic procedure; I_o is not defined as a behavior disposition.

If, on the other hand, "an IQ higher than 130" is taken as a theoretical term, the situation is entirely different. The same test procedure with S and R may still be accepted. But its specification is no longer regarded as an operational definition of the term. There cannot be a definition of the term on the basis of overt behavior. There may be various test procedures for the same concept. But no result of a single test nor of any number of tests is ever absolutely conclusive, although they may, under favorable circumstances, yield a high probability. Any statement ascribing the term in question to a person on the basis of a given test result may later be corrected in view of new evidence, even if there is no doubt that the test rules S were fulfilled and that the response R was made. If a psychologist accepts this non-conclusive, probabilistic character of a test, as, I suppose, practically all would do, then the concept tested cannot be a pure disposition and is best reconstructed as a theoretical term.

I think that, even on a prescientific level, many people would regard their psychological judgments about other people as in principle always

open to correction in view of later observations of their behavior. To the extent that someone is willing to change his judgments in this way, his use of psychological terms might be regarded as a beginning of the development which leads finally to theoretical terms. By the way, it would be interesting to make an empirical investigation of the degree of rigidity and flexibility shown by non-psychologists (including philosophers) in making and changing psychological statements about other people and about themselves. This would give a clearer indication of the nature of their concepts than any answers to direct questions about the concepts.

The distinction between intervening variables and theoretical constructs, often discussed since the article by MacCorquodale and Meehl, seems essentially the same or closely related to our distinction between pure dispositions and theoretical terms. "Theoretical construct" means certainly the same as here "theoretical term", *viz.*, a term which cannot be explicitly defined even in an extended observation language, but is introduced by postulates and not completely interpreted. The intervening variables are said to serve merely for a more convenient formulation of empirical laws and to be such that they can always be eliminated. Therefore it seems that they would be definable in a language similar to our extended observation language L'_O but containing also quantitative terms; thus they seem essentially similar to pure dispositions.

Among empiricists, it was especially Feigl who early recognized and continually emphasized the importance of theoretical laws (which he called "existential hypotheses"; see his (8)). And he showed in particular that in the present phase of psychology the use of theoretical concepts and laws constitutes one of the most important methodological problems and tasks. He made important contributions to the clarification of this problem, especially in his article (10); there he points out the close analogy with the earlier development of physics.

Psychological theories with theoretical terms will no doubt be further developed, probably to a much larger extent than so far. There are good reasons for expecting that a development of this kind will prove to be very fruitful, while without it the possible forms of theory construction are too limited to give a good chance for essential progress. This does not imply that the so-called "molar" approach in terms of observable behavior is to be rejected; on the contrary, this approach will always be an essential part of psychological investigation. What is wrong is only the principle which demands a restriction of the psychological method to this approach. The molar approach in psychology has a function similar to that of macrophysics both in the historical development and in present research. In all fields, the study of macro-events is the natural approach in the beginning; it leads to the first explanations of facts by the discovery of general regularities among observable properties ("empirical laws"); and it remains always indispensable as the source of confirming evidence for theories.

In physics great progress was made only by the construction of theories referring to unobservable events and micro-entities (atoms, electrons, etc.). Then it became possible to formulate a relatively small number of fundamental laws as postulates from which many empirical laws, both those already known and new ones, could be derived with the help of suitably constructed correspondence rules. In psychology analogous developments have begun from two different starting points. The one development began with the introspective approach. It proceeded from introspectively observed events (feelings, perceptions, images, beliefs, remembrances, etc.) to unconscious, i.e., introspectively not observable, events. These were first conceived as analogous to the observable events, e.g., unconscious feelings, beliefs, etc. Later also, new kinds of entities were introduced, e.g., drives, complexes, the id, the ego, and the like; however, the laws involving these entities are so far only stated in a qualitative form, which limits their explanatory and still more their predictive power. The other development began with the molar behavioristic approach. It started with a study of observable events of behavior, and then proceeded to dispositions, tendencies, abilities, potentialities for such events, and further to more abstract entities. Here the stage of the first quantitative laws has been reached.

Both these approaches in psychology will probably later converge toward theories of the central nervous system formulated in physiological terms. In this physiological phase of psychology, which has already begun, a more and more prominent role will be given to quantitative concepts and laws referring to micro-states described in terms of cells, molecules, atoms, fields, etc. And finally, micro-physiology may be based on micro-physics. This possibility of constructing finally all of science, including psychology, on the basis of physics, so that all theoretical terms are definable by those of physics and all laws derivable

from those of physics, is asserted by the thesis of *physicalism* (in its strong sense). (My recent views on the question of physicalism are not yet represented in my publications. Feigl (11) explains them, describes the historical development of physicalism in our movement, and gives an illuminating discussion of the theses of physicalism and the arguments for them.) By far the greater part of the development of psychology just outlined is, of course, today no more than a program for the future. Views vary a great deal as to the probability and even the possibility of such a development; and many will especially oppose, with either scientific or metaphysical arguments, the possibility of the last step, the assertion of physicalism. My personal impression, in view of the progress made within the last decades in psychology, physiology, the chemistry of complex organic molecules, and certain parts of physics, especially the theory of electronic computers, is that the whole development of psychology from the molar phase through the theoretical, the physiological, and the micro-physiological phases to the final foundation in micro-physics seems today much more probable and much less remote in time than it appeared even thirty years ago.

REFERENCES

1. Bridgman, P. W. *The Logic of Modern Physics.* New York: Macmillan, 1927.
2. Bridgman, P. W. *The Nature of Physical Theory.* Princeton: Princeton Univ. Pr., 1936.
3. Bridgman, P. W. "Operational Analysis," *Philosophy of Science,* 5:114–31 (1938).
4. Bridgman, P. W. "The Nature of Some of Our Physical Concepts," *British Journal for the Philosophy of Science,* 1:257–72 (February 1951); 2:25–44 (May 1951); 2:142–60 (August 1951). Reprinted as a separate monograph by Philosophical Lib., New York, 1952.
5. Carnap, Rudolf. "Testability and Meaning," *Philosophy of Science,* 3:420–68 (1936); 4:1–40 (1937). Reprinted as monograph by Whitlock's, Inc., New Haven, Connecticut, 1950. Parts reprinted in H. Feigl and M. Brodbeck (eds.), *Readings in the Philosophy of Science.* New York: Appleton-Century-Crofts, 1953.
6. Carnap, Rudolf. *Foundations of Logic and Mathematics,* Vol. I, No. 3 of the *International Encyclopedia of Unified Science.* Chicago: Univ. of Chicago Pr., 1939. A part ("The Interpretation of Physics") is reprinted in H. Feigl and M. Brodbeck (eds.), *Readings in the Philosophy of Science,* pp. 309–18. New York: Appleton-Century-Crofts, 1953.
7. Feigl, Herbert. "Operationism and Scientific Method," *Psychological Review,* 52:250–59 (1945). Reprinted, with some alterations, in H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis,* pp. 498–509. New York: Appleton-Century-Crofts, 1949.
8. Feigl, Herbert. "Existential Hypotheses: Realistic Vs. Phenomenalistic Interpretations," *Philosophy of Science,* 17:35–62 (1950).