VOL. LIV. No. 213.] [January, 1945.

MIND
A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—STUDIES IN THE LOGIC OF CONFIRMATION (I.).

To the memory of my wife, Eva Ahrends Hempel.

BY CARL G. HEMPEL.

1. *Objective of the Study.*[1]—The defining characteristic of an empirical statement is its capability of being tested by a confrontation with experimental findings, *i.e.* with the results of suitable experiments or " focussed " observations. This feature distinguishes statements which have empirical content both from the statements of the formal sciences, logic and mathematics, which require no experiential test for their validation, and from

[1] The present analysis of confirmation was to a large extent suggested and stimulated by a co-operative study of certain more general problems which were raised by Dr. Paul Oppenheim, and which I have been investigating with him for several years. These problems concern the form and the function of scientific laws and the comparative methodology of the different branches of empirical science. The discussion with Mr. Oppenheim of these issues suggested to me the central problem of the present essay. The more comprehensive problems just referred to will be dealt with by Mr. Oppenheim in a publication which he is now preparing.

In my occupation with the logical aspects of confirmation, I have benefited greatly by discussions with several students of logic, including Professor R. Carnap, Professor A. Tarski, and particularly Dr. Nelson Goodman, to whom I am indebted for several valuable suggestions which will be indicated subsequently.

A detailed exposition of the more technical aspects of the analysis of confirmation presented in this article is included in my article " A Purely Syntactical Definition of Confirmation ", *The Journal of Symbolic Logic*, vol. 8 (1943).

the formulations of transempirical metaphysics, which do not admit of any.

The testability here referred to has to be understood in the comprehensive sense of "testability in principle"; there are many empirical statements which, for practical reasons, cannot be actually tested at present. To call a statement of this kind testable in principle means that it is possible to state just what experiential findings, if they were actually obtained, would constitute favourable evidence for it, and what findings or "data", as we shall say for brevity, would constitute unfavourable evidence; in other words, a statement is called testable in principle, if it is possible to describe the kind of data which would confirm or disconfirm it.

The concepts of confirmation and of disconfirmation as here understood are clearly more comprehensive than those of conclusive verification and falsification. Thus, *e.g.* no finite amount of experiential evidence can conclusively verify a hypothesis expressing a general law such as the law of gravitation, which covers an infinity of potential instances, many of which belong either to the as yet inaccessible future, or to the irretrievable past; but a finite set of relevant data may well be "in accord with" the hypothesis and thus constitute confirming evidence for it. Similarly, an existential hypothesis, asserting, say, the existence of an as yet unknown chemical element with certain specified characteristics, cannot be conclusively proved false by a finite amount of evidence which fails to "bear out" the hypothesis; but such unfavourable data may, under certain conditions, be considered as weakening the hypothesis in question, or as constituting disconfirming evidence for it.[1]

While, in the practice of scientific research, judgments as to the confirming or disconfirming character of experiential data obtained in the test of a hypothesis are often made without hesitation and with a wide consensus of opinion, it can hardly be said that these judgments are based on an explicit theory providing general criteria of confirmation and of disconfirmation. In this respect, the situation is comparable to the manner in which deductive inferences are carried out in the practice of scientific research: This, too, is often done without reference to an explicitly stated system of rules of logical inference. But while criteria of valid deduction can be and have been supplied by

[1] This point as well as the possibility of conclusive verification and conclusive falsification will be discussed in some detail in section 10 of the present paper.

formal logic, no satisfactory theory providing general criteria of confirmation and disconfirmation appears to be available so far.

In the present essay, an attempt will be made to provide the elements of a theory of this kind. After a brief survey of the significance and the present status of the problem, I propose to present a detailed critical analysis of some common conceptions of confirmation and disconfirmation and then to construct explicit definitions for these concepts and to formulate some basic principles of what might be called the logic of confirmation.

2. *Significance and Present Status of the Problem.*—The establishment of a general theory of confirmation may well be regarded as one of the most urgent desiderata of the present methodology of empirical science.[1] Indeed, it seems that a precise analysis of the concept of confirmation is a necessary condition for an adequate solution of various fundamental problems concerning the logical structure of scientific procedure. Let us briefly survey the most outstanding of these problems.

(*a*) In the discussion of scientific method, the concept of relevant evidence plays an important part. And while certain " inductivist " accounts of scientific procedure seem to assume that relevant evidence, or relevant data, can be collected in the context of an inquiry prior to the formulation of any hypothesis, it should be clear upon brief reflection that relevance is a relative concept ; experiential data can be said to be relevant or irrelevant only with respect to a given hypothesis ; and it is the hypothesis which determines what kind of data or evidence are relevant for it. Indeed, an empirical finding is relevant for a hypothesis if and only if it constitutes either favourable or unfavourable evidence for it ; in other words, if it either confirms or disconfirms the hypothesis. Thus, a precise definition of relevance presupposes an analysis of confirmation and disconfirmation.

(*b*) A closely related concept is that of instance of a hypothesis. The so-called method of inductive inference is usually presented as proceeding from specific cases to a general hypothesis of which each of the special cases is an " instance " in the sense that it " conforms to " the general hypothesis in question, and thus constitutes confirming evidence for it.

Thus, any discussion of induction which refers to the establishment of general hypotheses on the strength of particular instances is fraught with all those logical difficulties—soon to be expounded

[1] Or of the " logic of science ", as understood by R. Carnap ; *cf. The Logical Syntax of Language* (New York and London, 1937), sect. 72, and the supplementary remarks in *Introduction to Semantics* (Cambridge, Mass., 1942), p. 250.

—which beset the concept of confirmation. A precise analysis of this concept is, therefore, a necessary condition for a clear statement of the issues involved in the problem complex of induction and of the ideas suggested for their solution—no matter what their theoretical merits or demerits may be.

(*c*) Another issue customarily connected with the study of scientific method is the quest for " rules of induction ". Generally speaking, such rules would enable us to " infer ", from a given set of data, that hypothesis or generalization which accounts best for all the particular data in the given set. Recent logical analyses have made it increasingly clear that this way of conceiving the problem involves a misconception: While the process of invention by which scientific discoveries are made is as a rule *psychologically guided and stimulated* by antecedent knowledge of specific facts, its results are *not logically determined* by them; the way in which scientific hypotheses or theories are discovered cannot be mirrored in a set of general rules of inductive inference.[1] One of the crucial considerations which lead to this conclusion is the following: Take a scientific theory such as the atomic theory of matter. The evidence on which it rests may be described in terms referring to directly observable phenomena, namely to certain " macroscopic " aspects of the various experimental and observational data which are relevant to the theory. On the other hand, the theory itself contains a large number of highly abstract, non-observational terms such as " atom ", " electron ", " nucleus ", " dissociation ", " valence " and others, none of which figures in the description of the observational data. An adequate rule of induction would therefore have to provide, for this and for every conceivable other case, mechanically applicable criteria determining unambiguously, and without any reliance on the inventiveness or additional scientific knowledge of its user, all those new abstract concepts which need to be created for the formulation of the theory that will account for the given evidence. Clearly, this requirement cannot be satisfied by any set of rules, however ingeniously devised; there can be no general rules of induction in the above sense; the demand for them rests on a confusion of logical and psychological issues. What determines the soundness of a hypothesis is not the way it

[1] See the lucid presentation of this point in Karl Popper's *Logik der Forschung* (Wien, 1935), esp. sect. 1, 2, 3, and 25, 26, 27; *cf.* also Albert Einstein's remarks in his lecture *On the Method of Theoretical Physics* (Oxford, 1933,) pp. 11 and 12. Also of interest in this context is the critical discussion of induction by H. Feigl in " The Logical Character of the Principle of Induction, " *Philosophy of Science*, vol. 1 (1934).

is arrived at (it may even have been suggested by a dream or a hallucination), but the way it stands up when tested, *i.e.* when confronted with relevant observational data. Accordingly, the quest for rules of induction in the original sense of canons of scientific discovery has to be replaced, in the logic of science, by the quest for general objective criteria determining (A) whether, and—if possible—even (B) to what degree, a hypothesis *H* may be said to be corroborated by a given body of evidence *E*. This approach differs essentially from the inductivist conception of the problem in that it presupposes not only *E*, but also *H* as given and then seeks to determine a certain logical relationship between them. The two parts of this latter problem can be restated in somewhat more precise terms as follows :

(A) To give precise definitions of the two non-quantitative relational concepts of confirmation and of disconfirmation ; *i.e.* to define the meaning of the phrases " *E* confirms *H* " and " *E* disconfirms *H* ". (When *E* neither confirms nor disconfirms *H*, we shall say that *E* is neutral, or irrelevant, with respect to *H*.)

(B) (1) To lay down criteria defining a metrical concept " degree of confirmation of *H* with respect to *E* ", whose values are real numbers ; or, failing this,

(2) To lay down criteria defining two relational concepts, " more highly confirmed than " and " equally well confirmed with ", which make possible a non-metrical comparison of hypotheses (each with a body of evidence assigned to it) with respect to the extent of their confirmation.

Interestingly, problem B has received much more attention in methodological research than problem A ; in particular, the various theories of the " probability of hypotheses " may be regarded as concerning this problem complex ; we have here adopted [1] the more neutral term " degree of confirmation " instead of " probability " because the latter is used in science in a definite technical sense involving reference to the relative frequency of the occurrence of a given event in a sequence, and it is at least an open question whether the degree of confirmation of a hypothesis can generally be defined as a probability in this statistical sense.

The theories dealing with the probability of hypotheses fall into two main groups : the " logical " theories construe probability as a logical relation between sentences (or propositions ;

[1] Following R. Carnap's usage in Testability and Meaning, *Philosophy of Science*, vols. 3 (1936) and 4 (1937) ; esp. sect. 3 (in vol. 3).

it is not always clear which is meant)[1]; the "statistical" theories interpret the probability of a hypothesis in substance as the limit of the relative frequency of its confirming instances among all relevant cases.[2] Now it is a remarkable fact that none of the theories of the first type which have been developed so far provides an explicit general definition of the probability (or degree of confirmation) of a hypothesis H with respect to a body of evidence E; they all limit themselves essentially to the construction of an uninterpreted postulational system of logical probability. For this reason, these theories fail to provide a complete solution of problem B. The statistical approach, on the other hand, would, if successful, provide an explicit numerical definition of the degree of confirmation of a hypothesis; this definition would be formulated in terms of the numbers of confirming and disconfirming instances for H which constitute the body of evidence E. Thus, a necessary condition for an adequate interpretation of degrees of confirmation as statistical probabilities is the establishment of precise criteria of confirmation and disconfirmation, in other words, the solution of problem A.

However, despite their great ingenuity and suggestiveness, the attempts which have been made so far to formulate a precise statistical definition of the degree of confirmation of a hypothesis seem open to certain objections,[3] and several authors[4] have expressed doubts as to the possibility of defining the degree of confirmation of a hypothesis as a metrical magnitude, though

[1] This group includes the work of such writers as Janina Hosiasson-Lindenbaum (*cf.* for instance, her article "Induction et analogie: Comparàison de leur fondement", MIND, vol. L (1941); also see p. 21, n. 2), H. Jeffreys, J. M. Keynes, B. O. Koopman, J. Nicod (see p. 9, n. 2), St. Mazurkiewicz, F. Waismann. For a brief discussion of this conception of probability, see Ernest Nagel, *Principles of the Theory of Probability* (Internat. Encyclopedia of Unified Science, vol. i, no. 6, Chicago, 1939), esp. sects. 6 and 8.

[2] The chief proponent of this view is Hans Reichenbach; *cf.* especially Ueber Induktion und Wahrscheinlichkeit, *Erkenntnis*, vol. v (1935), and *Experience and Prediction* (Chicago, 1938), Ch. V.

[3] *Cf.* Karl Popper, *Logik der Forschung* (Wien, 1935), sect. 80; Ernest Nagel, *l.c.*, sect. 8, and "Probability and the Theory of Knowledge", *Philosophy of Science*, vol. 6 (1939); C. G. Hempel, "Le problème de la vérité", *Theoria* (Göteborg), vol. 3 (1937), sect. 5, and "On the Logical Form of Probability Statements", *Erkenntnis*, vol. 7 (1937-38), esp. sect. 5. *Cf.* also Morton White, "Probability and Confirmation", *The Journal of Philosophy*, vol. 36 (1939).

[4] See, for example, J. M. Keynes, *A Treatise on Probability*, London, 1929, esp. Ch. III; Ernest Nagel, *Principles of the Theory of Probability* (*cf.* n. 1 above), esp. p. 70; compare also the somewhat less definitely sceptical statement by Carnap, *l.c.* (see p. 5, n. 1), sect. 3, p. 427.

some of them consider it as possible, under certain conditions, to solve at least the less exacting problem B (2), *i.e.* to establish standards of non-metrical comparison between hypotheses with respect to the extent of their confirmation. An adequate comparison of this kind might have to take into account a variety of different factors [1]; but again the numbers of the confirming and of the disconfirming instances which the given evidence includes will be among the most important of those factors.

Thus, of the two problems, A and B, the former appears to be the more basic one, first, because it does not presuppose the possibility of defining numerical degrees of confirmation or of comparing different hypotheses as to the extent of their confirmation; and second because our considerations indicate that any attempt to solve problem B—unless it is to remain in the stage of an axiomatized system without interpretation—is likely to require a precise definition of the concepts of confirming and disconfirming instance of a hypothesis before it can proceed to define numerical degrees of confirmation, or to lay down non-metrical standards of comparison.

(*d*) It is now clear that an analysis of confirmation is of fundamental importance also for the study of the central problem of what is customarily called epistemology; this problem may be characterized as the elaboration of "standards of rational belief" or of criteria of warranted assertibility. In the methodology of empirical science this problem is usually phrased as concerning the rules governing the test and the subsequent acceptance or rejection of empirical hypotheses on the basis of experimental or observational findings, while in its "epistemological" version the issue is often formulated as concerning the validation of beliefs by reference to perceptions, sense data, or the like. But no matter how the final empirical evidence is construed and in what terms it is accordingly expressed, the theoretical problem remains the same: to characterize, in precise and general terms, the conditions under which a body of evidence can be said to confirm, or to disconfirm, a hypothesis of empirical character; and that is again our problem A.

(*e*) The same problem arises when one attempts to give a precise statement of the empiricist and operationalist criteria for the empirical meaningfulness of a sentence; these criteria, as is well known, are formulated by reference to the theoretical

[1] See especially the survey of such factors given by Ernest Nagel in *Principles of the Theory of Probability* (*cf.* p. 6, n. 1), pp. 66-73.

testability of the sentence by means of experiential evidence [1]; and the concept of theoretical testability, as was pointed out earlier, is closely related to the concepts of confirmation and disconfirmation.[2]

Considering the great importance of the concept of confirmation, it is surprising that no systematic theory of the non-quantitative relation of confirmation seems to have been developed so far. Perhaps this fact reflects the tacit assumption that the concepts of confirmation and of disconfirmation have a sufficiently clear meaning to make explicit definitions unnecessary or at least comparatively trivial. And indeed, as will be shown below, there are certain features which are rather generally associated with the intuitive notion of confirming evidence, and which, at first, seem well suited to serve as defining characteristics of confirmation. Closer examination will reveal the definitions thus obtainable to be seriously deficient and will make it clear that an adequate definition of confirmation involves considerable difficulties.

Now the very existence of such difficulties suggests the question whether the problem we are considering does not rest on a false assumption: Perhaps there are no objective criteria of confirmation; perhaps the decision as to whether a given hypothesis is acceptable in the light of a given body of evidence is no more subject to rational, objective rules than is the process of inventing a scientific hypothesis or theory; perhaps, in the last analysis, it is a "sense of evidence", or a feeling of plausibility in view of the relevant data, which ultimately decides whether a hypothesis is scientifically acceptable.[3] This view is comparable to the opinion that the validity of a mathematical proof or of a logical argument has to be judged ultimately by reference to a feeling of soundness or convincingness; and both theses have to be rejected on analogous grounds: They involve a confusion of logical and psychological considerations. Clearly, the occurrence

[1] *Cf.*, for example, A. J. Ayer, *Language, Truth and Logic*, London and New York, 1936, Ch. I; R. Carnap, "Testability and Meaning" (*cf.* p. 5, n. 1) sects. 1, 2, 3; H. Feigl, *Logical Empiricism* (in *Twentieth Century Philosophy*, ed. by Dagobert D. Runes, New York, 1943); P. W. Bridgman, *The Logic of Modern Physics*, New York, 1928.

[2] It should be noted, however, that in his essay "Testability and Meaning" (*cf.* p. 5, n. 1) R. Carnap has constructed definitions of testability and confirmability which avoid reference to the concept of confirming and of disconfirming evidence; in fact, no proposal for the definition of these latter concepts is made in that study.

[3] A view of this kind has been expressed, for example, by M. Mandelbaum in "Causal Analyses in History", *Journal of the History of Ideas*, vol. 3 (1942); *cf.* esp. pp. 46-47.

or non-occurrence of a feeling of conviction upon the presentation of grounds for an assertion is a subjective matter which varies from person to person, and with the same person in the course of time ; it is often deceptive, and can certainly serve neither as a necessary nor as a sufficient condition for the soundness of the given assertion.[1] A rational reconstruction of the standards of scientific validation cannot, therefore, involve reference to a sense of evidence ; it has to be based on objective criteria. In fact, it seems reasonable to require that the criteria of empirical confirmation, besides being objective in character, should contain no reference to the specific subject-matter of the hypothesis or of the evidence in question ; it ought to be possible, one feels, to set up purely formal criteria of confirmation in a manner similar to that in which deductive logic provides purely formal criteria for the validity of deductive inferences.

With this goal in mind, we now turn to a study of the non-quantitative concept of confirmation. We shall begin by examining some current conceptions of confirmation and exhibiting their logical and methodological inadequacies ; in the course of this analysis, we shall develop a set of conditions for the adequacy of any proposed definition of confirmation ; and finally, we shall construct a definition of confirmation which satisfies those general standards of adequacy.

3. *Nicod's Criterion of Confirmation and its Shortcomings.*—We consider first a conception of confirmation which underlies many recent studies of induction and of scientific method. A very explicit statement of this conception has been given by Jean Nicod in the following passage : " Consider the formula or the law : *A entails B.* How can a particular proposition, or more briefly, a fact, affect its probability ? If this fact consists of the presence of B in a case of A, it is favourable to the law '*A entails B*' ; on the contrary, if it consists of the absence of B in a case of A, it is unfavourable to this law. It is conceivable that we have here the only two direct modes in which a fact can influence the probability of a law. . . . Thus, the entire influence of particular truths or facts on the probability of universal propositions or laws would operate by means of these two elementary relations which we shall call *confirmation* and *invalidation*." [2] Note that the applicability of this criterion is restricted to hypotheses of

[1] See Karl Popper's pertinent statement, *l.c.*, sect. 8.

[2] Jean Nicod, *Foundations of Geometry and Induction* (transl. by P. P. Wiener), London, 1930 ; p. 219 ; *cf.* also R. M. Eaton's discussion of " Confirmation and Infirmation ", which is based on Nicod's views ; it is included in Ch. III of his *General Logic*, New York, 1931.

the form "*A entails B*". Any hypothesis H of this kind may be expressed in the notation of symbolic logic [1] by means of a universal conditional sentence, such as, in the simplest case,

$$(x)(P(x) \supset Q(x)),$$

i.e. "For any object x: if x is a P, then x is a Q," or also "Occurrence of the quality P entails occurrence of the quality Q." According to the above criterion this hypothesis is confirmed by an object a, if a is P and Q; and the hypothesis is disconfirmed by a if a is P, but not Q. In other words, an object confirms a universal conditional hypothesis if and only if it satisfies both the antecedent (here: '$P(x)$') and the consequent (here: '$Q(x)$') of the conditional; it disconfirms the hypothesis if and only if it satisfies the antecedent, but not the consequent of the conditional; and (we add this to Nicod's statement) it is neutral, or irrelevant, with respect to the hypothesis if it does not satisfy the antecedent.

This criterion can readily be extended so as to be applicable also to universal conditionals containing more than one quantifier, such as "Twins always resemble each other", or, in symbolic notation, '$(x)(y)(\text{Twins}(x, y) \supset \text{Rsbl}(x, y))$'. In these cases, a confirming instance consists of an ordered couple, or triple, etc., of objects satisfying the antecedent and the consequent of the conditional. (In the case of the last illustration, any two persons who are twins and resemble each other would confirm the hypothesis; twins who do not resemble each other would disconfirm it; and any two persons not twins—no matter whether they resemble each other or not—would constitute irrelevant evidence.)

We shall refer to this criterion as Nicod's criterion.[2] It states explicitly what is perhaps the most common tacit interpretation of the concept of confirmation. While seemingly quite adequate, it suffers from serious shortcomings, as will now be shown.

(*a*) First, the applicability of this criterion is restricted to hypotheses of universal conditional form; it provides no standards of confirmation for existential hypotheses (such as "There exists organic life on other stars", or "Poliomyelitis is caused by some virus") or for hypotheses whose explicit formulation calls for the use of both universal and existential quantifiers (such as

[1] In this paper, only the most elementary devices of this notation are used; the symbolism is essentially that of *Principia Mathematica*, except that parentheses are used instead of dots, and that existential quantification is symbolized by '(E)' instead of by the inverted 'E'.

[2] This term is chosen for convenience, and in view of the above explicit formulation given by Nicod; it is not, of course, intended to imply that this conception of confirmation originated with Nicod.

"Every human being dies some finite number of years after his birth", or the psychological hypothesis, "You can fool all of the people some of the time and some of the people all of the time, but you cannot fool all of the people all of the time", which may be symbolized by '$(x)(Et)\mathrm{Fl}(x, t) . (Ex)(t)\mathrm{Fl}(x, t) . \sim (x)(t)\mathrm{Fl}(x, t)$', (where '$\mathrm{Fl}(x, t)$' stands for "You can fool (person) x at time t"). We note, therefore, the desideratum of establishing a criterion of confirmation which is applicable to hypotheses of any form.[1]

(*b*) We now turn to a second shortcoming of Nicod's criterion. Consider the two sentences

$$S_1 : \text{`}(x)(\mathrm{Raven}(x) \supset \mathrm{Black}(x))\text{'};$$
$$S_2 : \text{`}(x)(\sim \mathrm{Black}(x) \supset \sim \mathrm{Raven}(x))\text{'}$$

(*i.e.* "All ravens are black" and "Whatever is not black is not a raven"), and let a, b, c, d be four objects such that a is a raven and black, b a raven but not black, c not a raven but black, and d neither a raven nor black. Then, according to Nicod's criterion, a would confirm S_1, but be neutral with respect to S_2; b would disconfirm both S_1 and S_2; c would be neutral with respect to both S_1 and S_2, and d would confirm S_2, but be neutral with respect to S_1.

But S_1 and S_2 are logically equivalent; they have the same content, they are different formulations of the same hypothesis. And yet, by Nicod's criterion, either of the objects a and d would be confirming for one of the two sentences, but neutral with respect to the other. This means that Nicod's criterion makes confirmation depend not only on the content of the hypothesis, but also on its formulation.[2]

One remarkable consequence of this situation is that every hypothesis to which the criterion is applicable—*i.e.* every universal conditional—can be stated in a form for which there cannot possibly exist any confirming instances. Thus, *e.g.* the sentence

$$(x)[(\mathrm{Raven}(x) . \sim \mathrm{Black}(x)) \supset (\mathrm{Raven}(x) . \sim \mathrm{Raven}(x)]$$

is readily recognized as equivalent to both S_1 and S_2 above; yet no object whatever can confirm this sentence, *i.e.* satisfy both

[1] For a rigorous formulation of the problem, it is necessary first to lay down assumptions as to the means of expression and the logical structure of the language in which the hypotheses are supposed to be formulated; the desideratum then calls for a definition of confirmation applicable to any hypothesis which can be expressed in the given language. Generally speaking, the problem becomes increasingly difficult with increasing richness and complexity of the assumed "language of science".

[2] This difficulty was pointed out, in substance, in my article "Le problème de la vérité", *Theoria* (Göteborg), vol. 3 (1937), esp. p. 222.

its antecedent and its consequent; for the consequent is contradictory. An analogous transformation is, of course, applicable to any other sentence of universal conditional form.

4. *The Equivalence Condition.*—The results just obtained call attention to a condition which an adequately defined concept of confirmation should satisfy, and in the light of which Nicod's criterion has to be rejected as inadequate: *Equivalence condition:* Whatever confirms (disconfirms) one of two equivalent sentences, also confirms (disconfirms) the other.

Fulfilment of this condition makes the confirmation of a hypothesis independent of the way in which it is formulated; and no doubt it will be conceded that this is a necessary condition for the adequacy of any proposed criterion of confirmation. Otherwise, the question as to whether certain data confirm a given hypothesis would have to be answered by saying: "That depends on which of the different equivalent formulations of the hypothesis is considered"—which appears absurd. Furthermore—and this is a more important point than an appeal to a feeling of absurdity—an adequate definition of confirmation will have to do justice to the way in which empirical hypotheses function in theoretical scientific contexts such as explanations and predictions; but when hypotheses are used for purposes of explanation or prediction,[1] they serve as premisses in a deductive argument whose conclusion is a description of the event to be explained or predicted. The deduction is governed by the principles of formal logic, and according to the latter, a deduction which is valid will remain so if some or all of the premisses are replaced by different, but equivalent statements; and indeed, a scientist will feel free, in any theoretical reasoning involving certain hypotheses, to use the latter in whichever of their equivalent formulations is most convenient for the development of his conclusions. But if we adopted a concept of confirmation which did not satisfy the equivalence condition, then it would be possible, and indeed necessary, to argue in certain cases that it was sound scientific procedure to base a prediction on a given hypothesis if formulated in a sentence S_1, because a good deal of confirming evidence had

[1] For a more detailed account of the logical structure of scientific explanation and prediction, *cf.* C. G. Hempel, "The Function of General Laws in History", *The Journal of Philosophy*, vol. 39 (1942), esp. sects. 2, 3, 4. The characterization, given in that paper as well as in the above text, of explanations and predictions as arguments of a deductive logical structure, embodies an over-simplification: as will be shown in sect. 7 of the present essay, explanations and predictions often involve "quasi-inductive" steps besides deductive ones. This point, however, does not affect the validity of the above argument.

been found for S_1; but that it was altogether inadmissible to base the prediction (say, for convenience of deduction) on an equivalent formulation S_2, because no confirming evidence for S_2 was available. Thus, the equivalence condition has to be regarded as a necessary condition for the adequacy of any definition of confirmation.

5. *The "Paradoxes" of Confirmation.*—Perhaps we seem to have been labouring the obvious in stressing the necessity of satisfying the equivalence condition. This impression is likely to vanish upon consideration of certain consequences which derive from a combination of the equivalence condition with a most natural and plausible assumption concerning a sufficient condition of confirmation.

The essence of the criticism we have levelled so far against Nicod's criterion is that it certainly cannot serve as a necessary condition of confirmation; thus, in the illustration given in the beginning of section 3, the object a confirms S_1 and should therefore also be considered as confirming S_2, while according to Nicod's criterion it is not. Satisfaction of the latter is therefore not a necessary condition for confirming evidence.

On the other hand, Nicod's criterion might still be considered as stating a particularly obvious and important sufficient condition of confirmation. And indeed, if we restrict ourselves to universal conditional hypotheses in one variable [1]—such as S_1

[1] This restriction is essential: In its general form, which applies to universal conditionals in any number of variables, Nicod's criterion cannot even be construed as expressing a sufficient condition of confirmation. This is shown by the following rather surprising example: Consider the hypothesis S_1: $(x)(y)[\sim (R(x, y)) \supset (R(x, y) \,.\, \sim R(y, x))]$.

Let a, b be two objects such that $R(a, b)$ and $\sim R(b, a)$. Then clearly, the couple (a, b) satisfies both the antecedent and the consequent of the universal conditional S_1; hence, if Nicod's criterion in its general form is accepted as stating a sufficient condition of confirmation, (a, b) constitutes confirming evidence for S_1. However, S_1 can be shown to be equivalent to

$$S_2: (x)(y)R(x, y)$$

Now, by hypothesis, we have $\sim R(b, a)$; and this flatly contradicts S_2 and thus S_1. Thus, the couple (a, b), although satisfying both the antecedent and the consequent of the universal conditional S_1 actually constitutes disconfirming evidence of the strongest kind (conclusively disconfirming evidence, as we shall say later) for that sentence. This illustration reveals a striking and—as far as I am aware—hitherto unnoticed weakness of that conception of confirmation which underlies Nicod's criterion. In order to realize the bearing of our illustration upon Nicod's original formulation, let A and B be $\sim (R(x, y) \,.\, R(y, x))$ and $R(x, y) \,.\, \sim R(y, x)$ respectively. Then S_1 asserts that A entails B, and the couple (a, b) is a case of the presence of B in the presence of A; this should, according to Nicod, be favourable to S_1.

and S_2 in the above illustration—then it seems perfectly reasonable to qualify an object as confirming such a hypothesis if it satisfies both its antecedent and its consequent. The plausibility of this view will be further corroborated in the course of our subsequent analyses.

Thus, we shall agree that if a is both a raven and black, then a certainly confirms S_1: '(x) (Raven(x) $\supset$ Black(x))', and if d is neither black nor a raven, d certainly confirms S_2:

$$\text{'}(x)\ (\sim \text{Black}(x) \supset \sim \text{Raven}(x))\text{.'}$$

Let us now combine this simple stipulation with the equivalence condition: Since S_1 and S_2 are equivalent, d is confirming also for S_1; and thus, we have to recognize as confirming for S_1 any object which is neither black nor a raven. Consequently, any red pencil, any green leaf, and yellow cow, etc., becomes confirming evidence for the hypothesis that all ravens are black. This surprising consequence of two very adequate assumptions (the equivalence condition and the above sufficient condition of confirmation) can be further expanded: The following sentence can readily be shown to be equivalent to S_1: S_3: '(x) [(Raven(x) v $\sim$ Raven(x)) $\supset$ ($\sim$ Raven(x) v Black(x))]', *i.e.* "Anything which is or is not a raven is either no raven or black". According to the above sufficient condition, S_3 is certainly confirmed by any object, say e, such that (1) e is or is not a raven and, in addition, (2) e is not a raven or also black. Since (1) is analytic, these conditions reduce to (2). By virtue of the equivalence condition, we have therefore to consider as confirming for S_1 any object which is either no raven or also black (in other words: any object which is no raven at all, or a black raven).

Of the four objects characterized in section 3, a, c and d would therefore constitute confirming evidence for S_1, while b would be disconfirming for S_1. This implies that any non-raven represents confirming evidence for the hypothesis that all ravens are black.

We shall refer to these implications of the equivalence criterion and of the above sufficient condition of confirmation as the *paradoxes of confirmation.*

How are these paradoxes to be dealt with? Renouncing the equivalence condition would not represent an acceptable solution, as is shown by the considerations presented in section 4. Nor does it seem possible to dispense with the stipulation that an object satisfying two conditions, C_1 and C_2, should be considered as confirming a general hypothesis to the effect that any object which satisfies C_1, also satisfies C_2.

But the deduction of the above paradoxical results rests on one other assumption which is usually taken for granted, namely, that the meaning of general empirical hypotheses, such as that all ravens are black, or that all sodium salts burn yellow, can be adequately expressed by means of sentences of universal conditional form, such as '(x) (Raven$(x) \supset$ Black(x))' and '(x) (Sod. Salt$(x) \supset$ Burn Yellow (x))', etc. Perhaps this customary mode of presentation has to be modified; and perhaps such a modification would automatically remove the paradoxes of confirmation? If this is not so, there seems to be only one alternative left, namely to show that the impression of the paradoxical character of those consequences is due to misunderstanding and can be dispelled, so that no theoretical difficulty remains. We shall now consider these two possibilities in turn: The sub-sections 5.11 and 5.12 are devoted to a discussion of two different proposals for a modified representation of general hypotheses; in subsection 5.2, we shall discuss the second alternative, *i.e.* the possibility of tracing the impression of paradoxicality back to a misunderstanding.

5.11. It has often been pointed out that while Aristotelian logic, in agreement with prevalent every day usage, confers "existential import" upon sentences of the form "All *P*'s are *Q*'s", a universal conditional sentence, in the sense of modern logic, has no existential import; thus, the sentence

$$\text{`}(x)\ (\text{Mermaid}(x) \supset \text{Green}(x))\text{'}$$

does not imply the existence of mermaids; it merely asserts that any object either is not a mermaid at all, or a green mermaid; and it is true simply because of the fact that there are no mermaids. General laws and hypotheses in science, however—so it might be argued—are meant to have existential import; and one might attempt to express the latter by supplementing the customary universal conditional by an existential clause. Thus, the hypothesis that all ravens are black would be expressed by means of the sentence S_1: '(x) (Raven$(x) \supset$ Black(x)) . (Ex)Raven(x); and the hypothesis that no non-black things are ravens by S_2: '$(x)(\sim$ Black$(x) \supset \sim$ Raven(x)) . $(Ex) \sim$ Black(x). Clearly, these sentences are not equivalent, and of the four objects a, b, c, d characterized in section 3, part (b), only a might reasonably be said to confirm S_1, and only d to confirm S_2. Yet this method of avoiding the paradoxes of confirmation is open to serious objections:

(a) First of all, the representation of every general hypothesis by a conjunction of a universal conditional and an existential sentence would invalidate many logical inferences which are generally accepted as permissible in a theoretical argument. Thus, for example, the assertions that all sodium salts burn yellow, and that whatever does not burn yellow is no sodium salt are logically equivalent according to customary understanding and usage; and their representation by universal conditionals preserves this equivalence; but if existential clauses are added, the two assertions are no longer equivalent, as is illustrated above by the analogous case of S_1 and S_2.

(*b*) Second, the customary formulation of general hypotheses in empirical science clearly does not contain an existential clause, nor does it, as a rule, even indirectly determine such a clause unambiguously. Thus, consider the hypothesis that if a person after receiving an injection of a certain test substance has a positive skin reaction, he has diphtheria. Should we construe the existential clause here as referring to persons, to persons receiving the injection, or to persons who, upon receiving the injection, show a positive skin reaction? A more or less arbitrary decision has to be made; each of the possible decisions gives a different interpretation to the hypothesis, and none of them seems to be really implied by the latter.

(*c*) Finally, many universal hypotheses cannot be said to imply an existential clause at all. Thus, it may happen that from a certain astrophysical theory a universal hypothesis is deduced concerning the character of the phenomena which would take place under certain specified extreme conditions. A hypothesis of this kind need not (and, as a rule, does not) imply that such extreme conditions ever were or will be realized; it has no existential import. Or consider a biological hypothesis to the effect that whenever man and ape are crossed, the offspring will have such and such characteristics. This is a general hypothesis; it might be contemplated as a mere conjecture, or as a consequence of a broader genetic theory, other implications of which may already have been tested with positive results; but unquestionably the hypothesis does not imply an existential clause asserting that the contemplated kind of cross-breeding referred to will, at some time, actually take place.

While, therefore, the adjunction of an existential clause to the customary symbolization of a general hypothesis cannot be considered as an adequate *general* method of coping with the paradoxes of confirmation, there is a purpose which the use of an existential clause may serve very well, as was pointed out to me by Dr. Paul Oppenheim [1]: if somebody feels that objects of the

[1] This observation is related to Mr. Oppenheim's methodological studies referred to in p. 1, n. 1.

types *c* and *d* mentioned above are irrelevant rather than confirming for the hypothesis in question, and that qualifying them as confirming evidence does violence to the meaning of the hypothesis, then this may indicate that he is consciously or unconsciously construing the latter as having existential import; and this kind of understanding of general hypotheses is in fact very common. In this case, the "paradox" may be removed by pointing out that an adequate symbolization of the intended meaning requires the adjunction of an existential clause. The formulation thus obtained is more restrictive than the universal conditional alone; and while we have as yet set up no criteria of confirmation applicable to hypotheses of this more complex form, it is clear that according to every acceptable definition of confirmation objects of the types *c* and *d* will fail to qualify as confirming cases. In this manner, the use of an existential clause may prove helpful in distinguishing and rendering explicit different possible interpretations of a given general hypothesis which is stated in non-symbolic terms.

5.12. Perhaps the impression of the paradoxical character of the cases discussed in the beginning of section 5 may be said to grow out of the feeling that the hypothesis that all ravens are black is about ravens, and not about non-black things, nor about all things. The use of an existential clause was one attempt at expressing this presumed peculiarity of the hypothesis. The attempt has failed, and if we wish to reflect the point in question, we shall have to look for a stronger device. The idea suggests itself of representing a general hypothesis by the customary universal conditional, supplemented by the indication of the specific "field of application" of the hypothesis; thus, we might represent the hypothesis that all ravens are black by the sentence '(x) (Raven$(x) \supset$ Black(x))' (or any one of its equivalents), plus the indication "Class of ravens" characterizing the field of application; and we might then require that every confirming instance should belong to the field of application. This procedure would exclude the objects *c* and *d* from those constituting confirming evidence and would thus avoid those undesirable consequences of the existential-clause device which were pointed out in 5.11 (*c*). But apart from this advantage, the second method is open to objections similar to those which apply to the first: (*a*) The way in which general hypotheses are used in science never involves the statement of a field of application; and the choice of the latter in a symbolic formulation of a given hypothesis thus introduces again a considerable measure of arbitrariness. In particular, for a scientific hypothesis to the effect that all P's are Q's, the field of application cannot simply be said to be the class of all P's ; for a hypothesis such as that all sodium salts burn yellow finds important applications in tests with negative results ; *i.e.* it may be applied to a substance of which it is not known whether it contains sodium salts, nor whether it burns yellow ; and if the flame does not turn yellow, the hypothesis serves to establish the absence of sodium salts. The same is true of all other hypotheses used for tests of this type. (*b*) Again, the consistent use of a domain of application in the formulation of general hypotheses would involve considerable logical complications, and yet would have no counterpart in the theoretical procedure of science, where hypotheses are subjected to various kinds of logical transformation and inference without any consideration that might be regarded as referring to changes in the fields of application. This method of meeting the paradoxes would therefore amount to dodging the problem by means of an *ad hoc* device which cannot be justified by reference to actual scientific procedure.

5.2. We have examined two alternatives to the customary method of representing general hypotheses by means of universal conditionals ; neither of them proved an adequate means of precluding the paradoxes of confirmation. We shall now try to show that what is wrong does not lie in the customary way of construing and representing general hypotheses, but rather in our reliance on a misleading intuition in the matter : The impression of a paradoxical situation is not objectively founded ; it is a psychological illusion.

(*a*) One source of misunderstanding is the view, referred to before, that a hypothesis of the simple form " Every P is a Q " such as " All sodium salts burn yellow ", asserts something about a certain limited class of objects only, namely, the class of all P's. This idea involves a confusion of logical and practical considerations : Our interest in the hypothesis may be focussed upon its applicability to that particular class of objects, but the hypothesis nevertheless asserts something about, and indeed imposes restrictions upon, *all* objects (within the logical type of the variable occurring in the hypothesis, which in the case of our last illustration might be the class of all physical objects). Indeed, a hypothesis of the form " Every P is a Q " forbids the occurrence of any objects having the property P but lacking the property Q ; *i.e.* it restricts all objects whatsoever to the class of those which either lack the property P or also have the property Q. Now, every object either belongs to this class or falls outside it, and thus, every object—and not only the P's—either conforms to the hypothesis or violates it; there is no object which is not implicitly "referred to" by a hypothesis of this type. In particular, every object which either is no sodium salt or burns yellow conforms to, and thus "bears out" the hypothesis that all sodium salts burn yellow; every other object violates that hypothesis.

The weakness of the idea under consideration is evidenced also by the observation that the class of objects about which a hypothesis is supposed to assert something is in no way clearly determined, and that it changes with the context, as was shown in 5.12 (*a*).

(*b*) A second important source of the appearance of paradoxicality in certain cases of confirmation is exhibited by the following consideration.

Suppose that in support of the assertion "All sodium salts burn yellow" somebody were to adduce an experiment in which a piece of pure ice was held into a colourless flame and did not turn the flame yellow. This result would confirm the assertion, "Whatever does not burn yellow is no sodium salt", and consequently, by virtue of the equivalence condition, it would confirm the original formulation. Why does this impress us as paradoxical? The reason becomes clear when we compare the previous situation with the case of an experiment where an object whose chemical constitution is as yet unknown to us is held into a flame and fails to turn it yellow, and where subsequent analysis reveals it to contain no sodium salt. This outcome, we should no doubt agree, is what was to be expected on the basis of the hypothesis that all sodium salts burn yellow—no matter in which of its various equivalent formulations it may be expressed; thus, the data here obtained constitute confirming evidence for the hypothesis. Now the only difference between the two situations here considered is that in the first case we are told beforehand the test substance is ice, and we happen to "know anyhow" that ice contains no sodium salt; this has the consequence that the outcome of the flame-colour test becomes entirely irrelevant for the confirmation of the hypothesis and thus can yield no new evidence for us. Indeed, if the flame should not turn yellow, the hypothesis requires that the substance contain no sodium salt—and we know beforehand that ice does not—and if the flame should turn yellow, the hypothesis would impose no further restrictions on the substance; hence, either of the possible outcomes of the experiment would be in accord with the hypothesis.

The analysis of this example illustrates a general point: In the seemingly paradoxical cases of confirmation, we are often not actually judging the relation of the given evidence, E alone to the hypothesis H (we fail to observe the "methodological fiction", characteristic of every case of confirmation, that we have no relevant evidence for H other than that included in E); instead, we tacitly introduce a comparison of H with a body of evidence which consists of E in conjunction with an additional amount of information which we happen to have at our disposal; in our illustration, this information includes the knowledge (1) that the substance used in the experiment is ice, and (2) that ice contains no sodium salt. If we assume this additional information as given, then, of course, the outcome of the experiment can add no strength to the hypothesis under consideration. But if we are careful to avoid this tacit reference to additional knowledge (which entirely changes the character of the problem), and if we formulate the question as to the confirming character of the evidence in a manner adequate to the concept of confirmation as used in this paper, we have to ask: Given some object a (it happens to be a piece of ice, but this fact is not included in the evidence), and given the fact that a does not turn the flame yellow and is no sodium salt—does a then constitute confirming evidence for the hypothesis? And now—no matter whether a is ice or some other substance—it is clear that the answer has to be in the affirmative; and the paradoxes vanish.

So far, in section (*b*), we have considered mainly that type of paradoxical case which is illustrated by the assertion that any non-black non-raven constitutes confirming evidence for the hypothesis, "All ravens are black." However, the general idea just outlined applies as well to the even more extreme cases exemplified by the assertion that any non-raven as well as any black object confirms the hypothesis in question. Let us illustrate this by reference to the latter case. If the given evidence E—*i.e.* in the sense of the required methodological fiction, all our data relevant for the hypothesis—consists only of one object which, in addition, is black, then E may reasonably be said to support even the hypothesis that all objects are black, and *a fortiori* E supports the weaker assertion that all ravens are black. In this case, again, our factual knowledge that not all objects are black tends to create an impression of paradoxicality which is not justified on logical grounds. Other "paradoxical" cases of confirmation may be dealt with analogously, and it thus turns out that the "paradoxes of confirmation", as formulated above, are due to a misguided intuition in the matter rather than to a logical flaw in the two stipulations from which the "paradoxes" were derived.[1], [2]

[1] The basic idea of sect. (*b*) in the above analysis of the "paradoxes of confirmation" is due to Dr. Nelson Goodman, to whom I wish to reiterate my thanks for the help he rendered me, through many discussions, in clarifying my ideas on this point.

[2] The considerations presented in section (*b*) above are also influenced by, though not identical in content with, the very illuminating discussion of the "paradoxes" by the Polish methodologist and logician Janina Hosiasson-Lindenbaum; *cf.* her article "On Confirmation", *The Journal of Symbolic Logic*, vol. 5 (1940), especially sect. 4. Dr. Hosiasson's attention had been called to the paradoxes by the article referred to in p. 11, n. 2, and by discussions with the author. To my knowledge, hers has so far been the only publication which presents an explicit attempt to solve the problem. Her solution is based on a theory of degrees of confirmation, which is developed in the form of an uninterpreted axiomatic system (*cf.* also p. 6, n. 1, and part (*b*) in sect. 1 of the present article), and most of her arguments presuppose that theoretical framework. I have profited, however, by some of Miss Hosiasson's more general observations which proved relevant for the analysis of the paradoxes of the non-gradated relation of confirmation which forms the object of the present study.

One point in those of Miss Hosiasson's comments which rest on her theory of degrees of confirmation is of particular interest, and I should like to discuss it briefly. Stated in reference to the raven-hypothesis, it consists in the suggestion that the finding of one non-black object which is no raven, while constituting confirming evidence for the hypothesis, would increase the degree of confirmation of the hypothesis by a smaller amount than the finding of one raven which is black. This is said to be so because the class of all ravens is much less numerous than that of all non-black objects, so that—to put the idea in suggestive though somewhat misleading terms—the finding of one black raven confirms a larger portion of the total content of the hypothesis than the finding of one non-black non-raven. In fact, from the basic assumptions of her theory, Miss Hosiasson is able to derive a theorem according to which the above statement about the relative increase in degree of confirmation will hold provided that actually the number of all ravens is small compared with the number of all non-black objects. But is this last numerical assumption actually warranted in the present case and analogously in all other "paradoxical" cases? The answer depends in part upon the logical structure of the language of science. If a "co-ordinate language" is used, in which, say, finite space-time regions figure as individuals, then the raven-hypothesis assumes some such form as "Every space-time region which contains a raven, contains something black"; and even if the total number of ravens ever to exist is finite, the class of space-time regions containing a raven has the power of the continuum, and so does the class of space-time regions containing something non-black; thus, for a co-ordinate language of the type under consideration, the above numerical assumption is not warranted. Now the use of a co-ordinate language may appear quite artificial in this particular illustration; but it will seem very appropriate in many other contexts, such as, *e.g.*, that of physical field theories. On the other hand, Miss Hosiasson's numerical assumption may well be justified on the basis of a "thing language", in which physical objects of finite size function

6. *Confirmation Construed as a Relation between Sentences.*—Our analysis of Nicod's criterion has so far led to two main results: The rejection of that criterion in view of several deficiencies, and the emergence of the equivalence condition as a necessary condition of adequacy for any proposed definition of confirmation. Another aspect of Nicod's criterion requires consideration now. In our formulation of the criterion, confirmation was construed as a dyadic relation between an object or an ordered set of objects, representing the evidence, and a sentence, representing the hypothesis. This means that confirmation was conceived of as a semantical relation [1] obtaining between certain extra-linguistic objects [2] on one hand and certain sentences on the other. It is possible, however, to construe confirmation in an alternative fashion as a relation between two sentences, one describing the given evidence, the other expressing the hypothesis. Thus, *e.g.* instead of saying that an object *a* which is both a raven and black (or the "fact" of *a* being both a raven and black) confirms the hypothesis, "All ravens are black", we may say that the evidence sentence, "*a* is a raven, and *a* is black", confirms the hypothesis-sentence (briefly, the hypothesis), "All ravens are black". We shall adopt this conception of confirmation as a relation between sentences here for the following reasons: First, the evidence adduced in support or criticism of a scientific hypothesis is always expressed in sentences, which frequently have the character of observation reports; and second, it will prove very fruitful to pursue the parallel, alluded to in section 2 above, between the concepts of confirmation and of logical consequence. And just as in the theory of the consequence relation, *i.e.* in deductive logic, the premisses of which a given conclusion is a consequence are construed as sentences rather than as "facts", so we propose to construe the data which confirm a given hypothesis as given in the form of sentences.

The preceding reference to observation reports suggests a certain restriction which might be imposed on evidence sentences. Indeed, the evidence adduced in support of a scientific hypothesis

as individuals. Of course, even on this basis, it remains an empirical question, for every hypothesis of the form "All *P*'s are *Q*'s", whether actually the class of non-*Q*'s is much more numerous than the class of *P*'s; and in many cases this question will be very difficult to decide.

[1] For a detailed account of this concept, see C. W. Morris, *Foundations of the Theory of Signs* (Internat. Encyclopedia of Unified Science, vol. i, no. 2, Chicago, 1938), and R. Carnap, *Introduction to Semantics* (Cambridge, Mass., 1942), esp. sects. 4 and 37.

[2] Instead of making the first term of the relation an object or a sequence of objects, we might construe it as a "state of affairs" (or perhaps as a "fact", or a "proposition", as Nicod puts it), such as that state of affairs which consists in *a* being a black raven, etc.

or theory consists, in the last analysis, in data accessible to what is loosely called " direct observation ", and such data are expressible in the form of " observation reports ". In view of this consideration, we shall restrict the evidence sentences which form the domain of the relation of confirmation, to sentences of the character of observation reports. In order to give a precise meaning to the concept of observation report, we shall assume that we are given a well-determined " language of science ", in terms of which all sentences under consideration, hypotheses as well as evidence sentences, are formulated. We shall further assume that this language contains, among other terms, a clearly delimited " observational vocabulary " which consists of terms designating more or less directly observable attributes of things or events, such as, say, " black ", " taller than ", " burning with a yellow light ", etc., but no theoretical constructs such as " aliphatic compound ", " circularly polarized light ", " heavy hydrogen ", etc.

We shall now understand by a hypothesis any sentence which can be expressed in the assumed language of science, no matter whether it is a generalized sentence, containing quantifiers, or a particular sentence referring only to a finite number of particular objects. An observation report will be construed as a finite class (or a conjunction of a finite number) of observation sentences ; and an observation sentence as a sentence which either asserts or denies that a given object has a certain observable property (such as " *a* is a raven ", " *d* is not black "), or that a given sequence of objects stand in a certain observable relation (such as " *a* is between *b* and *c* ").

Now the concept of observability itself obviously is relative to the techniques of observation used. What is unobservable to the unaided senses may well be observable by means of suitable devices such as telescopes, microscopes, polariscopes, lie-detectors, Gallup-polls, etc. If by direct observation we mean such observational procedures as do not make use of auxiliary devices, then such property terms as " black ", " hard ", " liquid ", " cool ", and such relation terms as " above ", " between ", " spatially coincident ", etc., might be said to refer to directly observable attributes ; if observability is construed in a broader sense, so as to allow for the use of certain specified instruments or other devices, the concept of observable attribute becomes more comprehensive. If, in our study of confirmation, we wanted to analyze the manner in which the hypotheses and theories of empirical science are ultimately supported by " evidence of the senses ", then we should have to require that observation reports refer exclusively to directly observable attributes. This view was taken, for simplicity and concreteness, in the preceding parts of this section. Actually, however, the general logical characteristics of that relation which obtains between a hypothesis and a group of empirical statements which " support " it, can be studied in isolation from this restriction to direct observability. All we will assume here is that in the context of the scientific test of a given hypothesis or theory, certain specified techniques of observation have been agreed upon; these determine an observational vocabulary, namely a set of terms designating properties and relations observable by means of the accepted techniques. For our purposes it is entirely sufficient that these terms, constituting the " observational vocabulary ", be given. An observation sentence is then defined simply as a sentence affirming or denying that a given object, or sequence of objects, possesses one of those observable attributes.[1]

Let it be noted that we do not require an observation sentence to be true, nor to be accepted on the basis of actual observations; rather, an observation sentence expresses something that is decidable by means of the accepted techniques of observation; in other words: An observation sentence describes a possible outcome of the accepted observational techniques; it asserts something that might conceivably be established by means of those

[1] The concept of observation sentence has, in the context of our study, a status and a logical function closely akin to that of the concepts of protocol statement or basis sentence, etc., as used in many recent studies of empiricism. However, the conception of observation sentence which is being proposed in the present study is more liberal in that it renders the discussion of the logical problems of testing and confirmation independent of various highly controversial epistemological issues; thus, *e.g.* we do not stipulate that observation reports must be about psychic acts, or about sense perceptions (*i.e.* that they have to be expressed in terms of a vocabulary of phenomenology, or of introspective psychology). According to the conception of observation sentence adopted in the present study, the " objects " referred to in an observation sentence may be construed in any one of the senses just referred to, or in various other ways; for example, they might be space-time regions, or again physical objects such as stones, trees, etc. (most of the illustrations given throughout this article represent observation sentences belonging to this kind of " thing-language "); all that we require is that the few very general conditions stated above be satisfied.

These conditions impose on observation sentences and on observation reports certain restrictions with respect to their form; in particular, neither kind of sentence may contain any quantifiers. This stipulation recommends itself for the purposes of the logical analysis here to be undertaken; but we do not wish to claim that this formal restriction is indispensable. On the contrary, it is quite possible and perhaps desirable also to allow for observation sentences containing quantifiers: our simplifying assumption is introduced mainly in order to avoid considerable logical complications in the definition of confirmation.

techniques. Possibly, the term "observation-type sentence" would be more suggestive; but for convenience we give preference to the shorter term. An analogous comment applies, of course, to our definition of an observation report as a class or a conjunction of observation sentences. The need for this broad conception of observation sentences and observation reports is readily recognized: Confirmation as here conceived is a logical relationship between sentences, just as logical consequence is. Now whether a sentence S_2 is a consequence of a sentence S_1 does not depend on whether S_1 is true (or known to be true), or not; and analogously, the criteria of whether a given statement, expressed in terms of the observational vocabulary, confirms a certain hypothesis cannot depend on whether the statements in the report are true, or based on actual experience, or the like. Our definition of confirmation must enable us to indicate what kind of evidence *would* confirm a given hypothesis *if* it were available; and clearly the sentence characterizing such evidence can be required only to express something that might be observed, but not necessarily something that has actually been established by obsrvation.

It may be helpful to carry the analogy between confirmation and consequence one step further. The truth or falsity of S_1 is irrelevant for the question of whether S_2 is a consequence of S_1 (whether S_2 can be validly inferred from S_1); but in a logical inference which justifies a sentence S_2 by showing that it is a logical consequence of a conjunction of premisses, S_1, we can be certain of the truth of S_2 only if we know S_1 to be true. Analogously, the question of whether an observation report stands in the relation of confirmation to a given hypothesis does not depend on whether the report states actual or fictitious observational findings; but for a decision as to the soundness or acceptability of a hypothesis which is confirmed by a certain report, it is of course necessary to know whether the report is based on actual experience or not. Just as a conclusion of a logical inference, in order to be reliably true must be (a1) validly inferred from (a2) a set of true premisses, so a hypothesis, in order to be scientifically acceptable, must be (b1) formally confirmed by (b2) reliable reports on observational findings.

The central problem of this essay is to establish general criteria for the formal relation of confirmation as referred to in (b1); the analysis of the concept of a reliable observation report, which belongs largely to the field of pragmatics,[1] falls outside the scope of the present study. One point, however, deserves mention here: A statement of the form of an observation report (for

[1] An account of the concept of pragmatics may be found in the publications listed in p. 22, n. 1.

example, about the position of the pointer of a certain thermograph at 3 a.m.) may be accepted or rejected in science either on the basis of direct observation, or because it is indirectly confirmed or disconfirmed by other accepted observation sentences (in the example, these might be sentences describing the curve traced by the pointer during the night), and because of this possibility of indirect confirmation, our study has a bearing also on the question of the acceptance of hypotheses which have themselves the form of observation reports.

The conception of confirmation as a relation between sentences analogous to that of logical consequence suggests yet another specification for the attempted definition of confirmation: While logical consequence has to be conceived of as a basically semantical relation between sentences, it has been possible, for certain languages, to establish criteria of logical consequence in purely syntactical terms.[1] Analogously, confirmation may be conceived of as a semantical relation between an observation report and a hypothesis; but the parallel with the consequence relation suggests that it should be possible, for certain languages, to establish purely syntactical criteria of confirmation. The subsequent considerations will indeed eventuate in a definition of confirmation based on the concept of logical consequence and other purely syntactical concepts.

The interpretation of confirmation as a logical relation between sentences involves no essential change in the central problem of the present study. In particular, all the points made in the preceding sections can readily be rephrased in accordance with this interpretation. Thus, for example, the assertion that an object a which is a swan and white confirms the hypothesis '(x) (Swan$(x) \supset$ White(x))' can be expressed by saying that the observation report 'Swan(a) . White(a)' confirms that hypothesis. Similarly, the equivalence condition can be reformulated as follows: If an observation report confirms a certain sentence, then it also confirms every sentence which is logically equivalent with the latter. Nicod's criterion as well as our grounds for rejecting it can be re-formulated along the same lines. We presented Nicod's concept of confirmation as referring to a relation between non-linguistic objects on one hand and sentences on the other because this approach seemed to approximate most closely Nicod's own formulations, and because it enabled us to avoid certain technicalities which are actually unnecessary in that context.

(*To be concluded*)

[1] *Cf.* especially the two publications by R. Carnap listed in p. 3, n. 1.