The Classical Approach

Carl G. Hempel

STUDIES IN THE LOGIC OF CONFIRMATION

To the memory of my wife,
Eva Ahrends 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 finding, 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 experimental test for their validation, and from the formulations of trans-empirical 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.[2]

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

Reprinted from *Mind* (1945) by permission of the editor of *Mind*.

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 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.[3] 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—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 complex problem 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.[4] One of the crucial considerations which lead to this conclusion is the following: Take a scientic 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, nonobservational 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 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 nonquantitative 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 nonmetrical 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[5] 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; it is not always clear which is meant)[6]; 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.[7] 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,[8] and several authors[9] have expressed doubts as to the possibility of defining the degree of confirmation of a hypothesis as a metrical magnitude, though 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 nonmetrical 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;[10] 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 nonmetrical 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 testability of the sentence by means of experimental evidence;[11] and the concept of theoretical testability, as was pointed out earlier, is closely related to the concepts of confirmation and disconfirmation.[12]

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.[13] 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 or nonoccurrence 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.[14] 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 for malcriteria 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*."[15] Note that the applicability of the criterion is restricted to hypotheses of the form "*A entails B*." Any hypothesis *H* of this kind may be expressed in the notation of symbolic logic[16] 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, "Occurrences 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 to this 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 who 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.[17] 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 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 "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 "$(x)(Et)\text{Fl}(x, t) \cdot (Ex)(t)\text{Fl}(x, t) \cdot \sim (x)(t)\text{Fl}(x, t)$," (where "$\text{Fl}(x, t)$" 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.[18]

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

$$S_1\text{: } (x)(\text{Raven}(x) \supset \text{Black}(x));$$

$$S_2\text{: } (x)(\sim\text{Black}(x) \supset \sim\text{Raven}(x))$$

(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.[19]

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:

$(x)[(\text{Raven}(x)\cdot\sim\text{Black}(x)) \supset (\text{Raven}(x)\cdot\sim\text{Raven}(x)]$

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 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,[20] 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 hypothesis 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 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[21]—such as 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: "$(x)(\text{Raven}(x) \supset \text{Black}(x))$," and if *d* is neither black nor a raven, *d* certainly confirms S_2:

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

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, any 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: $(x)[(\text{Raven}(x) \vee \sim \text{Raven}(x)) \supset (\sim\text{Raven}(x) \vee \text{Black}(x))]$, 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 consideration 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 "$(x)(\text{Raven}(x) \supset \text{Black}(x))$" and "$(x)(\text{Sod. Salt}(x) \supset \text{Burn Yellow}(x))$," 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 subsections 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:

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

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: "$(x)(\text{Raven}(x) \supset \text{Black}(x)) \cdot (Ex)\text{Raven}(x)$"; and the hypothesis that no non-black things are ravens by S_2: "$(x)(\sim\text{Black}(x) \supset \sim\text{Raven}(x)) \cdot (Ex) \sim\text{Black}(x))$." 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 crossbreeding 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:[22] If somebody feels that objects of the 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 nonsymbolic 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 "$(x)(\text{Raven}(x) \supset \text{Black}(x))$" (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.[23,24]

* * *

8. CONDITIONS OF ADEQUACY FOR ANY DEFINITION OF CONFIRMATION

The two most customary conceptions of confirmation, which were rendered explicit in Nicod's criterion and in the prediction criterion, have thus been found unsuitable for a general definition of confirmation. Besides this negative result, the preceding analysis has also exhibited certain logical characteristics of scientific prediction, explanation, and testing, and it has led to the establishment of certain standards which an adequate definition of confirmation has to satisfy. These standards include the equivalence condition and the requirement that the definition of confirmation be applicable to hypotheses of any degree of logical complexity, rather than to the simplest type of universal conditional only. An adequate definition of confirmation, however, has to satisfy several further logical requirements, to which we now turn.

First of all, it will be agreed that any sentence which is entailed by, i.e., a logical consequence of, a given observation report has to be considered as confirmed by that report: Entailment is a special case of confirmation. Thus, e.g., we want to say that the observation report "*a* is black" confirms the sentence (hypothesis) "a is black or grey"; and—to refer to one of the illustrations given in the preceding section—the observation sentence $R_2(a, b)$ should certainly be confirming evidence for the sentence $(Ez)R_2(a, z)$. We are therefore led to the stipulation that any adequate definition of confirmation must insure the fulfilment of the

(8.1) *Entailment condition.* Any sentence which is entailed by an observation report is confirmed by it.[25]

This condition is suggested by the preceding consideration, but of course not proved by it. To make it a standard of adequacy for the definition of confirmation means to lay down the stipulation that a proposed definition of confirmation will be rejected as logically inadequate if it is not constructed in such a way that (8.1) is unconditionally satisfied. An analogous remark applies to the subsequently proposed further standards of adequacy.

Second, an observation report which confirms certain hypotheses would invariably be qualified as confirming any consequence of those hypotheses. Indeed, any such consequence is but an assertion of all or part of the combined content of the original hypotheses and has therefore to be regarded as confirmed by any evidence which confirms the original hypotheses. This suggests the following condition of adequacy:

(8.2) *Consequence Condition.* If an observation report confirms every one of a class K of sentences, then it also confirms any sentence which is a logical consequence of K.

If (8.2) is satisfied, then the same is true of the following two more special conditions:

(8.21) *Special Consequence Condition.* If an observation report confirms a hypothesis H, then it also confirms every consequence of H.

(8.22) *Equivalence Condition.* If an observation report confirms a hypothesis H, then it also confirms every hypothesis which is logically equivalent with H.

(This follows from (8.21) in view of the fact that equivalent hypotheses are mutual consequences of each other.) Thus, the satisfaction of the consequence condition entails that of our earlier equivalence condition, and the latter loses its status of an independent requirement.

In view of the apparent obviousness of these conditions, it is interesting to note that the definition of confirmation in terms of successful prediction, while satisfying the equivalence condition, would violate the consequence condition. Consider, for example, the formulation of the prediction-criterion given in the earlier part of the preceding section. Clearly, if the observational findings B_2 can be predicted on the basis of the findings B_1 by means of the hypothesis H, the same prediction is obtainable by means of any equivalent hypothesis, but not generally by means of a weaker one.

On the other hand, any prediction obtainable by means of H can obviously also be established by means of any hypothesis which is stronger than H, i.e., which logically entails H. Thus, while the consequence condition stipulates in effect that whatever confirms a given hypothesis also confirms any weaker hypothesis, the relation of confirmation defined in terms of successful prediction would satisfy the condition that whatever confirms a given hypothesis also confirms every stronger one.

But is this "converse consequence condition," as it might be called, not reasonable enough, and should it not even be included among our standards of adequacy for the definition of confirmation? The second of these two suggestions can be readily disposed of: The adoption of the new condition, in addition to (8.1) and (8.2), would have the consequence that any observation report B would confirm any hypothesis H whatsoever. Thus, e.g., if B is the report "a is a raven" and H is Hooke's law, then, according to (8.1), B confirms the sentence "a is a raven," hence B would, according to the converse consequence condition, confirm the stronger sentence "a is a raven, and Hooke's law holds"; and finally, by virtue of (8.2), B would confirm H, which is a consequence of the last sentence. Obviously, the same type of argument can be applied in all other cases.

But is it not true, after all, that very often observational data which confirm a hypothesis H are considered also as confirming a stronger hypothesis? Is it not true, for example, that those experimental findings which confirm Galileo's law, or Kepler's laws, are considered also as confirming Newton's law of gravitation?[26] This is indeed the case, but this does not justify the acceptance of the converse entailment condition as a general rule of the logic of confirmation; for in the cases just mentioned, the weaker hypothesis is connected with the stronger one by a logical bond of a particular kind: It is essentially a substitution instance of the stronger one; thus, e.g., while the law of gravitation refers to the force obtaining between any two bodies, Galileo's law is a specialization referring to the case where one of the bodies is the earth, the other an object near its surface. In the preceding case, however, where Hooke's law was shown to be confirmed by the observation report that a is a raven, this situation does not prevail; and here, the rule that whatever confirms a given hypothesis also confirms any stronger one becomes an entirely absurd principle. Thus, the converse consequence condition does not provide a sound general condition of adequacy.[27]

A third condition remains to be stated:[28]

(8.3) *Consistency Condition.* Every logically consistent observation report is logically compatible with the class of all the hypotheses which it confirms.

The two most important implications of this requirement are the following:

(8.31) Unless an observation report is self-contradictory,[29] it does not confirm any hypothesis with which it is not logically compatible.

(8.32) Unless an observation report is self-contradictory, it does not confirm any hypotheses which contradict each other.

The first of these corollaries will readily be accepted; the second, however,—and consequently (8.3) itself—will perhaps be felt to embody a too severe restriction. It might be pointed out, for example, that a finite set of measurements concerning the variation of one physical magnitude, x, with another, y, may conform to, and thus be said to confirm, several different hypotheses as to the particular mathematical function in terms of which the relationship of x and y can be expressed; but such hypotheses are incompatible because to at least one value of x, they will assign different values of y.

No doubt it is possible to liberalize the formal standards of adequacy in line with these considerations. This would amount to dropping (8.3) and (8.32) and retaining only (8.31). One of the effects of this measure would be that when a logically consistent observation report B confirms each of two hypotheses, it does not necessarily confirm their conjunction; for the hypotheses might be mutually incompatible, hence their conjunction self-contradictory; consequently, by (8.31), B could not confirm it. This consequence is intuitively rather awkward, and one might therefore feel inclined to suggest that while (8.3) should be dropped and (8.31) retained, (8.32) should be replaced by the requirement (8.33): If an observation sentence confirms each of two hypotheses, then it also confirms their conjunction. But it can readily be shown that by virtue of (8.2) this set of conditions entails the fulfilment of (8.32).

If, therefore, the condition (8.3) appears to be too rigorous, the most obvious alternative would seem to lie in replacing (8.3) and its corollaries by the much weaker condition (8.31) alone; and it is an important problem whether an intuitively adequate definition of confirmation can be constructed which satisfies (8.1), (8.2), and (8.31), but not (8.3). One of the great advantages of a definition which satisfies (8.3) is that it sets a limit, so to speak, to the strength of the hypotheses which can be confirmed by given evidence.[30]

The remainder of the present study, therefore, will be concerned exclusively with the problem of establishing a definition of confirmation which satisfies the more severe formal conditions represented by (8.1), (8.2), and (8.3) together.

The fulfilment of these requirements, which may be regarded as general laws of the logic of confirmation, is of course only a necessary, not a sufficient, condition for the adequacy of any proposed definition of confirmation. Thus, e.g., if "*B* confirms *H*" were defined as meaning "*B* logically entails *H*," then the above three conditions would clearly be satisfied; but the definition would not be adequate because confirmation has to be a more comprehensive relation than entailment (the latter might be referred to as the special case of *conclusive* confirmation). Thus, a definition of confirmation, to be acceptable, also has to be materially adequate: It has to provide a reasonably close approximation to that conception of confirmation which is implicit in scientific procedure and methodological discussion. That conception is vague and to some extent quite unclear, as I have tried to show in earlier parts of this paper; therefore, it would be too much to expect full agreement as to the material adequacy of a proposed definition of confirmation; on the other hand, there will be rather general agreement on certain points; thus, e.g., the identification of confirmation with entailment, or the Nicod criterion of confirmation as analyzed above, or any definition of confirmation by reference to a "sense of evidence" will probably now be admitted not to be adequate approximations to that concept of confirmation which is relevant for the logic of science.

On the other hand, the soundness of the logical analysis (which, in a clear sense, always involves a logical reconstruction) of a theoretical concept cannot be gauged simply by our feelings of satisfaction at a certain proposed analysis; and if there are, say, two alternative proposals for defining a term on the basis of a logical analysis, and if both appear to come fairly close to the intended meaning, then the choice has to be made largely by reference to such features as the logical properties of the two reconstructions, and the comprehensiveness and simplicity of the theories to which they lead.

9. THE SATISFACTION CRITERION OF CONFIRMATION

As has been mentioned before, a precise definition of confirmation requires reference to some definite "language of science," in which all observation reports and all hypotheses under consideration are assumed to be formulated, and whose logical structure is supposed to be precisely determined. The more complex this language, and the richer its logical means of expression, the more difficult it will be, as a rule, to establish an adequate definition of confirmation for it. However, the problem has been solved at least for certain cases: With respect to languages of a comparatively simple logical structure, it has been possible to construct an explicit definition of confirmation which satisfies all of the above logical requirements, and which appears to be intuitively rather adequate. An exposition of the technical details of this definition has been published elsewhere;[31] in the present study, which is concerned with the general logical and methodological aspects of the problem of confirmation rather than with technical detail, it will be attempted to characterize the definition of confirmation thus obtained as clearly as possible with a minimum of technicalities.

Consider the simple case of the hypothesis H: $(x)(\text{Raven}(x) \supset \text{Black}(x))$, where "Raven" and "Black" are supposed to be terms of our observational vocabulary. Let B be an observation report to the effect that $\text{Raven}(a) \cdot \text{Black}(a) \cdot \sim \text{Raven}(c) \cdot \text{Black}(c) \cdot \sim \text{Raven}(d) \cdot \sim \text{Black}(d)$. Then B may be said to confirm H in the following sense: There are three objects altogether mentioned in B, namely a, c, and d; and as far as these are concerned, B informs us that all those which are ravens (i.e., just the object a) are also black.[32] In other words, from the information contained in B we can infer that the hypothesis H does hold true within the finite class of those objects which are mentioned in B.

Let us apply the same consideration to a hypothesis of a logically more complex structure. Let H be the hypothesis "Everybody likes somebody"; in symbols: $(x)(Ey)\text{Likes}(x, y)$, i.e. for every (person) x, there exists at least one (not necessarily different person) y such that x likes y. (Here again, "Likes" is supposed to be a relation-term which occurs in our observational vocabulary.) Suppose now that we are given an observation report B in which the names of two persons, say e and f, occur. Under what conditions shall we say that B confirms H? The previous illustration suggests the answer: If from B we can infer that H is satisfied within the finite class $\{e, f\}$; i.e. that within $\{e, f\}$ everybody likes somebody. This in turn means that e likes e or f, and f likes e or f. Thus, B would be said to confirm H if B entailed the statement "e likes e or f, and f likes e or f." This latter statement will be called the development of H for the finite class $\{e, f\}$.

The concept of *development of a hypothesis, H, for a finite class of individuals, C,* can be defined in a general fashion; the development of H for C states what H would assert if there existed exclusively those objects which are elements of C. Thus, e.g., the development of the hypothesis $H_1 = (x)(P(x) \vee Q(x))$ (i.e., "Every object has the property P or the property Q") for the class $\{a, b\}$ is $(P(a) \vee Q(a)) \cdot (P(b) \vee Q(b))$ (i.e., "a has the property P or the property Q, and b has the property P or the property Q"); the development of the existential hypothesis H_2 that at least one object has the property P, i.e., $(Ex)P(x)$, for $\{a, b\}$ is $P(a) \vee P(b)$; the development of a hypothesis which contains no quantifiers, such as H_3: $P(c) \vee Q(c)$ is defined as that hypothesis itself, no matter what the reference class of individuals is.

A more detailed formal analysis based on considerations of this type leads to the introduction of a general relation of confirmation in two steps; the first consists in defining a special relation of direct confirmation along the lines just indicated; the second step then defines the general relation of confirmation by reference to direct confirmation.

Omitting minor details, we may summarize the two definitions as follows:

(9.1 Df.) An observation report B directly confirms a hypothesis H if B entails the development of H for the class of those objects which are mentioned in B.

(9.2 Df.) An observation report B confirms a hypothesis H if H is entailed by a class of sentences each of which is directly confirmed by B.

The criterion expressed in these definitions might be called the satisfaction criterion of confirmation because its basic idea consists in construing a hypothesis as confirmed by a given observation report if the hypothesis is satisfied in the finite class of those individuals which are mentioned in the report. Let us now apply the two definitions to our last examples: The observation report B_1: $P(a) \cdot Q(b)$ directly confirms (and therefore also confirms) the hypothesis H_1, because it entails the development of H_1 for the class $\{a, b\}$, which was given above. The hypothesis H_3 is not directly confirmed by B, because its development, i.e., H_3 itself, obviously is not entailed by B_1. However, H_3 is entailed by H_1, which is directly confirmed by B_1; hence, by virtue of (9.2), B_1 confirms H_3.

Similarly, it can readily be seen that B_1 directly confirms H_2.

Finally, to refer to the first illustration given in this section: The observation report Raven(a) · Black(a) · ~ Raven(c) · ~ Black(c) · ~ Raven(d) · ~ Black(d) confirms (even directly) the hypothesis (x)(Raven(x) ⊃ Black(x)), for it entails the development of the latter for the class $\{a, c, d\}$, which can be written as follows: (Raven(a) ⊃ Black(a)) · (Raven(c) ⊃ Black(c)) · (Raven(d) ⊃ Black(d)).

It is now easy to define disconfirmation and neutrality:

(9.3 Df.) An observation report B disconfirms a hypothesis H if it confirms the denial of H.

(9.4 Df.) An observation report B is neutral wth respect to a hypothesis H if B neither confirms nor disconfirms H.

By virtue of the criteria laid down in (9.2), (9.3), (9.4), every consistent observation report, B, divides all possible hypotheses into three mutually exclusive classes: Those confirmed by B, those disconfirmed by B, and those with respect to which B is neutral.

The definition of confirmation here proposed can be shown to satisfy all the formal conditions of adequacy embodies in (8.1), (8.2), and (8.3) and their consequences; for the condition (8.2) this is easy to see; for the other conditions the proof of more complicated.[33]

Furthermore, the application of the above definition of confirmation is not restricted to hypotheses of universal conditional form (as Nicod's criterion is, for example), nor to universal hypotheses in general; it applies, in fact, to any hypothesis which can be expressed by means of property and relation terms of the observational vocabulary of the given language, individual names, the customary connective symbols for "not," "and," "or," "if-then," and any number of universal and existential quantifiers.

Finally, as is suggested by the preceding illustrations as well as by the general considerations which underlie the establishment of the above definition, it seems that we have obtained a definition of confirmation which also is materially adequate in the sense of being a reasonable approximation to the intended meaning of confirmation.

* * *

NOTES

1. The present analysis of confirmation was to a large extent suggested and stimulated by a cooperative 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. VIII (1943).

2. This point as well as the possibility of conclusive verification and conclusive falsification will be discussed in some detail in Sec. 10 of the present paper.

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

4. See the lucid presentation of this point in Karl Popper's *Logik der Forschung* (Wien, 1935), esp. Secs. 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. 11–12. 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. I (1934).

5. Following R. Carnap's usage in "Testability and Meaning," *Philosophy of Science*, Vols. III (1936) and IV (1937); esp. Sec. 3 (in Vol. III).

6. This group includes the work of such writers as Janina Hosiasson-Lindenbaum (cf. for instance, her article "Induction et analogie: Comparaison de leur fondement," *Mind*, Vol. L (1941); (also see n. 24), H. Jeffreys, J. M. Keynes, B. O. Koopman, J. Nicod (see n. 15), 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. Secs. 6 and 8.

7. 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), Chap. v.

8. Cf. Karl Popper, *Logik der Forschung* (Wien, 1935), Sec. 80; Ernest Nagel, *loc. cit.*, Sec. 8, and "Probability and the Theory of Knowledge," *Philosophy of Science*, Vol. VI (1939); C. G. Hempel, "Le problème de la vérité," *Theoria* (Göteborg), Vol. III (1937), Sec. 5, and "On the Logical Form of Probability Statements," *Erkenntnis*, Vol. VII (1937–38), esp. Sec. 5. Cf. also Morton White, "Probability and Confirmation," *The Journal of Philosophy*, Vol. XXXVI (1939).

9. See, for example, J. M. Keynes, *A Treatise on Probability* (London, 1929), esp. Chap. iii; Ernest Nagel, *Principles of the Theory of Probability* (cf. n. 6 above), esp. p. 70. Compare also the somewhat less definitely sceptical statement by Carnap, *loc. cit.* (see n. 5), Sec. 3, p. 427.

10. See especially the survey of such factors given by Ernest Nagel in *Principles of the Theory of Probability* (cf. n. 6), pp. 66–73.

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

12. It should be noted, however, that in his essay "Testability and Meaning" (cf. n. 5), 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.

13. 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. III (1942); cf. esp. pp. 46–47.

14. See Karl Popper's pertinent statement, *loc. cit.*, Sec. 8.

15. Jean Nicod, *Foundations of Geometry and Induction*, trans. 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 Chap. iii of his *General Logic* (New York, 1931).

16. 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*."

17. 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.

18. 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 hypotheses 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."

19. This difficulty was pointed out, in substance, in my article "Le problème de la vértité," *Taeoria* (Göteborg), Vol. III (1937), esp. p. 222.

20. 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. XXXIX (1942), esp. Secs. 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 oversimplification: as will be shown in Sec. 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.

21. 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:

$$(x)(y)[\sim(R(x, y)\cdot R(y, x)) \supset (R(x, y)\cdot\sim R(y, x))].$$

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

$$S_2\text{: } (x)(y)R(x, y)$$

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 $\sim(R(x,y) \cdot R(y,x))$ and $R(x, y) \cdot \sim R(y, x)$ 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.

22. This observation is related to Mr. Oppenheim's methodological studies referred to in n. 1.

23. 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.

24. 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. V (1940), especially Sec. 4. Dr. Hosiasson's attention had been called to the paradoxes by the article referred to in 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. n. 6 and part (b) in Sec. 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 "coordinate 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 coordinate language of the type under consideration, the above numerical assumption is not warranted. Now the use of a coordinate 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 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.

25. As a consequence of this stipulation, a contradictory observation report, such as {Black(a), ~Black(a)} confirms every sentence, because it has every sentence as a consequence. Of course, it is possible to exclude the possibility of contradictory observation reports altogether by a slight restriction of the definition of "observation report." There is, however, no important reason to do so.

26. Strictly speaking, Galileo's law and Kepler's laws can be deduced from the law of gravitation only if certain additional hypotheses—including the laws of motion—are presupposed; but this does not affect the point under discussion.

27. William Barrett, in a paper entitled "Discussion on Dewey's Logic" (*The Philosophical Review,* Vol. L (1941), pp. 305 ff., esp. p. 312) raises some questions closely related to what we have called above the consequence condition and the converse consequence condition. In fact, he invokes the latter (without stating it explicitly) in an argument which is designed to show that "not every observation which confirms a sentence need also confirm all its consequences," in other words, that the special consequence condition (8.21) need not always be satisfied. He supports his point by reference to "the simplest case: the sentence *C* is an abbreviation of *A* · *B*, and the observation 0 confirms *A, and so C,* but is irrelevant to *B*, which is a consequence of *C*." (Italics mine.)

For reasons contained in the above discussion of the consequence condition and the converse consequence condition, the application of the latter in the case under consideration seems to us unjustifiable, so that the illustration does not prove the author's point; and indeed, there seems to be every reason to preserve the unrestricted validity of the consequence condition. As a matter of fact, Mr. Barrett himself argues that "the degree of confirmation for the consequence of a sentence cannot be less than that of the sentence itself"; this is indeed quite sound; but it is hard to see how the recognition of this principle can be reconciled with a renunciation of the special consequence condition, since the latter may be considered simply as the correlate, for the non-gradated relation of confirmation, of the former principle which is adapted to the concept of degree of confirmation.

28. For a fourth condition, see n. 33.

29. A contradictory observation report confirms every hypothesis (cf. n. 8) and is, of course, incompatible with every one of the hypotheses it confirms.

30. This was pointed out to me by Dr. Nelson Goodman. The definition later to be outlined in this essay, which satisfies conditions (8.1), (8.2), and (8.3), lends itself, however, to certain generalizations which satisfy only the more liberal conditions of adequacy just considered.

31. In my article referred to in n. 1. The logical structure of the languages to which the definition in question is applicable is that of the lower functional calculus with individual constants, and with predicate constants of any degree. All sentences of the language are assumed to be formed exclusively by means of predicate constants, individual constants, individual variables, universal

and existential quantifiers for individual variables, and the connective symbols of denial, conjunction, alternation, and implication. The use of predicate variables or of the identity sign is not permitted.

As to the predicate constants, they are all assumed to belong to the observational vocabulary, i.e. to denote a property or a relation observable by means of the accepted techniques. ("Abstract" predicate terms are supposed to be defined in terms of those of the observational vocabulary and then actually to be replaced by their *definientia,* so that they never occur explicitly.)

As a consequence of these stipulations, an observation report can be characterized simply as a conjunction of sentences of the kind illustrated by $P(a)$, $\sim P(b)$, $R(c, d)$, $\sim R(e, f)$, etc., where P, R, etc., belong to the observational vocabulary, and a, b, c, d, e, f, etc., are individual names, denoting specific objects. It is also possible to define an observation report more liberally as any sentence containing no quantifiers, which means that besides conjunctions also alternations and implication sentences formed out of the above kind of components are included among the observation reports.

32. I am indebted to Dr. Nelson Goodman for having suggested this idea; it initiated all those considerations which finally led to the definition to be outlined below.

33. For these proofs, see the article referred to in Part I, n. 1. I should like to take this opportunity to point out and to remedy a certain defect of the definition of confirmation which was developed in that article, and which has been outlined above: This defect was brought to my attention by a discussion with Dr. Olaf Helmer.

It will be agreed that an acceptable definition of confirmation should satisfy the following further condition which might well have been included among the logical standards of adequacy set up in Sec. 8 above: (8.4). If B_1 and B_2 are logically equivalent observation reports and B_1 confirms (disconfirms, is neutral with respect to) a hypothesis H, then B_2, too, confirms (disconfirms, is neutral with respect to) H. This condition is indeed satisfied if observation reports are construed, as they have been in this article, as classes or conjunctions of observation sentences. As was indicated at the end of n. 14, however, this restriction of observation reports to a conjunctive form is not essential; in fact, it has been adopted here only for greater convenience of exposition, and all the preceding results, including especially the definitions and theorems of the present section, remain applicable without change if observation reports are given the more liberal interpretation characterized at the end of n. 14. (In this case, if P and Q belong to the observational vocabulary, such sentences as $P(a) \vee Q(a)$, $P(a) \vee Q(b)$, etc., would qualify as observation reports.) This broader conception of observation reports was therefore adopted in the article referred to in Part I, n. 1; but it has turned out that in this case, the definition of confirmation summarized above does not generally satisfy the requirement (8.4). Thus, e.g., the observation reports, $B_1 = P(a)$ and $B_2 = P(a) \cdot (Q(b) \vee \sim Q(b))$ are logically equivalent, but while B_1 confirms (and even directly confirms) the hypothesis $H_1 = (x)P(x)$, the second report does not do so, essentially because it does not entail $P(a) \cdot P(b)$, which is the development of H_1 for the class of those objects mentioned in B_2. This deficiency can be remedied as follows: The fact that B_2 fails to confirm H_1 is obviously due to the circumstance that B_2 contains the individual constant b, without asserting anything about b: The object b is mentioned only in an analytic component of B_2. The atomic constituent $Q(b)$ will therefore be said to occur (twice) inessentially in B_2. Generally, an atomic constituent A of a molecular sentence S will be said to occur inessentially in S if by virtue of the rules of the sentential calculus S is equivalent to a molecular sentence in which A does not occur at all. Now an object will be said to be mentioned inessentially in an observation report if it is mentioned only in such components of that report as occur inessentially in it. The sentential calculus clearly provides mechanical procedures for deciding whether a given observation report mentions any object inessentially, and for establishing equivalent formulations of the same report in which no object is mentioned inessentially. Finally, let us say that an object is mentioned essentially in an observation report if it is mentioned, but not only mentioned inessentially, in that report. Now we replace (9.1) by the following definition:

(9.1*a*) An observation report B directly confirms a hypothesis H if B entails the development of H for the class of those objects which are mentioned essentially in B.