Theory and Prediction in Science

Rudolf Carnap
Department of Philosophy, University of Chicago

The activity of the scientist is twofold: he observes events, and he constructs laws intended to explain the facts observed. These two activities are closely interconnected. We do not first accumulate isolated facts and only then start looking for laws to explain them. It is true, we consider the results of previous observations when we hypothetically set up a law. But on the other hand, in collecting our observational data and especially in choosing suitable experimental arrangements, we are guided by at least a rudimentary system of laws, even if it is merely a vague guess at a regularity or correlation.

Similar to the relation between observed facts and experimental laws is the relation between these laws and others of a wider generality and a higher degree of abstraction. The abstract laws are based upon the experimental laws; but in laying down an experimental law we take into consideration not only observational results but also more general laws already available or anticipated. Finally, we try to construct a theory as an integrated system of laws of different levels of generality and abstraction. The development of science proceeds simultaneously in both directions: more and more observational results are collected with regard to, and with the help of, theories; and theories are constructed and modified with regard to observed facts.

From the purely theoretical point of view, the construction of a *theory* is the end and goal of science. When we are interested, on the other hand, in the practical application of science, we ask for *predictions.* We want to obtain knowledge, or at least a reasonable conjecture, as to what events we may expect. This is indispensable as a basis for reasonable decisions in practical activities. However, predictions are necessary not only for the practical application but also for purely theoretical purposes. In order to test a theory we derive predictions with its help; then we observe whether and to what extent the facts bear out the predictions, and these results are taken as the basis for our judgment of the theory.

There are some important issues in the methodology of science receiving at present much discussion by scientists and philosophers which center around this problem of the relation between a theory and observed facts. All agree today in rejecting the view prevalent in the 19th Century that a law or theory is inferred from facts by induction. We may still speak of in

Opening remarks at the meeting of Section L, AAAS, St. Louis, 28 March 1946.

ductive inference, since it is customary to do so; but if we do, we should keep in mind that our meaning is not that of the 19th Century. We all agree (1) that there is not in inductive inference the certainty there is in deductive inference, and (2) that there are no fixed rules which lead, so to speak, mechanically from a given body of facts to a theory. The theory is not *discovered* by a wholly rational or regulated procedure; in addition to knowledge of the relevant facts and to experience in working with other theories, such nonrational factors as intuition or the inspiration of genius play a decisive role. Of course, once the theory is proposed, there may be a rational procedure for *examining* it. Thus it becomes clear that the relation between a theory and the observational evidence available is, strictly speaking, not that of *inferring* the one from the other but rather that of *judging* the theory on the basis of the evidence when both are given.

A similar problem exists with regard to the relation between the prediction of a future event and the observational evidence available at the present time. Here, likewise, the question is, strictly speaking, not one of *inferring* the prediction from the evidence bunt rather of *judging* the reliability or strength of the given prediction with respect to the given evidence. This problem is similar to that concerning a theory, but there are also some differences. Some believe that, with respect to the evidence at hand, our primary judgments concern the reliability of theories, and that judgments as to the reliability of predictions of single events are derivative, in the sense that they depend upon the reliability of the theories used in making the predictions. Others believe that judgments about predictions are primary, and that the reliability of a theory cannot mean anything else than the reliability of the predictions to which the theory leads. And according to a third view, there is a general concept of the reliability of an hypothesis of any form with respect to given evidence. Theories and singular predictions are in this case regarded as merely two special kinds of hypotheses.

Instead of the term 'reliability,' other terms are sometimes used with regard to either theories, or predictions—for instance, 'probability,' 'weight', 'strength,' 'degree of confirmation,' etc. I shall continue deliberately to use the vague and somewhat ambiguous term 'reliability' in order to leave open the still controversial question as to the nature of the concept here involved, which might be prejudged by the choice of one of the more specific terms.

Let me now briefly outline some of the problems concerning the task of judging, on the basis of given observational evidence, the reliability of an hypothesis, which may be either a theory, a prediction, or perhaps & statement of still another form. I shall briefly outline the controversial issues, since it is not possible to explain here the reasons for the various views.

The first question is whether there is any rational and objective procedure at all for judging the reliability of hypotheses—any procedure objective in the sense of not being dependent on personal whim or bias. If there are inductive judgments in this sense, then there arises the question: What is their form? They may be merely comparative, stating, for instance, that the available evidence gives stronger support to the one than to the other of two given, incompatible hypotheses, or that a given hypothesis is supported more by the one than by the other of two bodies of evidence. But can we go beyond mere comparison and make quantitative inductive judgments attributing a numerical value to the reliability of an hypothesis? This is one of the most controversial issues. Some believe that this is not possible (e.g. von Mises, *10*). Others, such as Keynes (*8*), Koopman (*9*), and Nagel (*11*), regard it as possible within a special, narrowly limited domain—for instance, far the prediction of a result in a game of chance. Still others, including Jeffreys (*7*), Carnap (*1*), and Helmer, Hempel, and Oppenheim (*5*, *6*), think that quantitative judgments are possible in a wide field, or even for hypotheses of any kind.

Thus, even the form of inductive judgments about the reliability of hypotheses is under debate. Still more controversial are the problems of the nature and meaning of these judgments. What precisely is meant by 'reliability,' 'probability,' or 'degree of confirmation'? By what method can we obtain inductive judgments? What is the foundation of their validity?

Some authors think that the concept of the reliability of an hypothesis (called 'probability' or 'weight') must be based upon, or even identified with, the statistical concept of probability—that is, relative frequency within an infinite sequence (Reichenbach, *14*). Therefore, an inductive judgment about an hypothesis is here regarded as empirical. Related to this view are the modern methods of mathematical statistics, which are likewise based upon the frequency conception of probability (especially those of R. A. Fisher, *3*, *4*, and Neyman and Pearson, *12*, *13*; for a brief survey see Wald, *15*). Here we find different forms of inductive judgments; the hypotheses to be judged are usually statistical hypotheses concerning the distribution of properties or of values of physical magnitudes within a given domain. The weakest judgments express merely the acceptance or rejection of the given hypothesis on the basis of the given evidence. Stronger methods have been developed which yield an estimate for the value of a parameter occurring in a given statistical hypothesis. Some statisticians apply a numerical concept of reliability to a special kind of statistical hypotheses (Fisher's 'fiducial probability').

Finally, there are students who regard the concept of the reliability of an hypothesis (called by some 'probability' and by others 'degree of confirmation') as a purely logical concept; therefore, the inductive judgments are here not empirical but analytic, similar to the deductive judgment saying that a given conclusion follows logically from given premises. Some adherents of this conception, such as Keynes (*8*), and Koopman (*9*), restrict the inductive judgments entirely or chiefly to the comparative form, while others, such as Jeffreys (*7*), admit quantitative judgments in a wider field. Recently, attempts have been made at an explicit definition of a quantitative concept of reliability ('degree of confirmation') applicable to hypotheses of any form expressible in a certain simple language (Carnap, *1*, and Helmer, Hempel, and Oppenheim, *5*, *6*).

All these various approaches to the problem of inductive judgments of theories, predictions, or other hypotheses are still in a state of development. In particular, the question as to the relations of the different methods to one another is still under debate. Sometimes adherents of one method regard the other methods as unsound. But there are others, and I am among them, who believe that the various approaches are not incompatible but supplementary to each other (2). Further clarification of the nature of each method and its relation to the other methods is required.

References

1. CARNAP, R. *Phil. Sci.*, 1945, 12, 72.
2. CARNAP, R. *Phil. phenom. Res.*, 1945, 5, 513.
3. FISHER, R. A. *Philos. Trans. roy. Soc. Lond.* (Series A), 1922, 222, 309.
4. FISHER, R. A. *The design of experiments.* Edinburgh-London: Oliver and Boyd, 1935, 1942.
5. HELMER, O., and OPPENHEIM, P. J. *symb. Logic*, 1945, 10, 25.
6. HEMPEL, C. G., and OPPENHEIM, P. *Phil. Sci.*, 1945. 12, 98.
7. JEFFREYS, H. *Theory of probability*. Oxford, Engl.: Clarendon Press, 1939.
8. KEYNES, J. M. *A treatise on probability.* London-New York : Macmillan, 1921, 1929.
9. KOOPMAN, B. O. *Bull. Amer. math. Soc.*, 1940, 46, 703.
10. MISES, R. VON. *Probability, statistics, and truth.* New York : Macmillan, 1939.
11. NAGEL, E. *Principles of the theory of probability.* Chicago: Univ. Chicago Press, 1939.
12. NEYMAN, J., and PEARSON, E. S. *Statistical research memoirs*. London:

Univ. College, 1930, 1938. (Vole. 1, I1.)

13. NEYMAN, J. *Lectures and conferences on mathematical statistics*. Washington, D. C.: Graduate School, U. S. Dept. of Agriculture, 1938.
14. REICHENBACH, H. *Wahrscheinlichkeitslehre*. Leiden: Sijthoff, 1935; *Experience and prediction*. Chicago: Univ. Chicago Press, 1938.
15. WALD, A. *On the principles of statistical inference*. South Bend, Ind.: Univ. Notre Dame, 1942.