STUDIES IN THE LOGIC OF EXPLANATION

CARL G. HEMPEL AND PAUL OPPENHEIM[1]

§1. *Introduction.*

To explain the phenomena in the world of our experience, to answer the question "why?" rather than only the question "what?", is one of the foremost objectives of all rational inquiry; and especially, scientific research in its various branches strives to go beyond a mere description of its subject matter by providing an explanation of the phenomena it investigates. While there is rather general agreement about this chief objective of science, there exists considerable difference of opinion as to the function and the essential characteristics of scientific explanation. In the present essay, an attempt will be made to shed some light on these issues by means of an elementary survey of the basic pattern of scientific explanation and a subsequent more rigorous analysis of the concept of law and of the logical structure of explanatory arguments.

The elementary survey is presented in Part I of this article; Part II contains an analysis of the concept of emergence; in Part III, an attempt is made to exhibit and to clarify in a more rigorous manner some of the peculiar and perplexing logical problems to which the familiar elementary analysis of explanation gives rise. Part IV, finally, is devoted to an examination of the idea of explanatory power of a theory; an explicit definition, and, based on it, a formal theory of this concept are developed for the case of a scientific language of simple logical structure.

PART I. ELEMENTARY SURVEY OF SCIENTIFIC EXPLANATION

§2. *Some illustrations.*

A mercury thermometer is rapidly immersed in hot water; there occurs a temporary drop of the mercury column, which is then followed by a swift rise. How is this phenomenon to be explained? The increase in temperature affects at first only the glass tube of the thermometer; it expands and thus provides a larger space for the mercury inside, whose surface therefore drops. As soon as by heat conduction the rise in temperature reaches the mercury, however, the latter expands, and as its coefficient of expansion is considerably larger than that of

[1] This paper represents the outcome of a series of discussions among the authors; their individual contributions cannot be separated in detail. The technical developments contained in Part IV, however, are due to the first author, who also put the article into its final form.

Some of the ideas presented in Part II were suggested by our common friend, Kurt Grelling, who, together with his wife, became a victim of Nazi terror during the war. Those ideas were developed by Grelling, in a discussion by correspondence with the present authors, of emergence and related concepts. By including at least some of that material, which is indicated in the text, in the present paper, we feel that we are realizing the hope expressed by Grelling that his contributions might not entirely fall into oblivion.

We wish to express our thanks to Dr. Rudolf Carnap, Dr. Herbert Feigl, Dr. Nelson Goodman, and Dr. W. V. Quine for stimulating discussions and constructive criticism.

glass, a rise of the mercury level results.—This account consists of statements of two kinds. Those of the first kind indicate certain conditions which are realized prior to, or at the same time as, the phenomenon to be explained; we shall refer to them briefly as antecedent conditions. In our illustration, the antecedent conditions include, among others, the fact that the thermometer consists of a glass tube which is partly filled with mercury, and that it is immersed into hot water. The statements of the second kind express certain general laws; in our case, these include the laws of the thermic expansion of mercury and of glass, and a statement about the small thermic conductivity of glass. The two sets of statements, if adequately and completely formulated, explain the phenomenon under consideration: They entail the consequence that the mercury will first drop, then rise. Thus, the event under discussion is explained by subsuming it under general laws, i.e., by showing that it occurred in accordance with those laws, by virtue of the realization of certain specified antecedent conditions.

Consider another illustration. To an observer in a row boat, that part of an oar which is under water appears to be bent upwards. The phenomenon is explained by means of general laws—mainly the law of refraction and the law that water is an optically denser medium than air—and by reference to certain antecedent conditions—especially the facts that part of the oar is in the water, part in the air, and that the oar is practically a straight piece of wood.—Thus, here again, the question "*Why* does the phenomenon happen?" is construed as meaning "according to what general laws, and by virtue of what antecedent conditions does the phenomenon occur?"

So far, we have considered exclusively the explanation of particular events occurring at a certain time and place. But the question "Why?" may be raised also in regard to general laws. Thus, in our last illustration, the question might be asked: Why does the propagation of light conform to the law of refraction? Classical physics answers in terms of the undulatory theory of light, i.e. by stating that the propagation of light is a wave phenomenon of a certain general type, and that all wave phenomena of that type satisfy the law of refraction. Thus, the explanation of a general regularity consists in subsuming it under another, more comprehensive regularity, under a more general law.—Similarly, the validity of Galileo's law for the free fall of bodies near the earth's surface can be explained by deducing it from a more comprehensive set of laws, namely Newton's laws of motion and his law of gravitation, together with some statements about particular facts, namely the mass and the radius of the earth.

§3. The basic pattern of scientific explanation.

From the preceding sample cases let us now abstract some general characteristics of scientific explanation. We divide an explanation into two major constituents, the explanandum and the explanans[2]. By the explanandum, we

[2] These two expressions, derived from the Latin *explanare*, were adopted in preference to the perhaps more customary terms "explicandum" and "explicans" in order to reserve the latter for use in the context of explication of meaning, or analysis. On explication in this sense, cf. Carnap, [Concepts], p. 513.—Abbreviated titles in brackets refer to the bibliography at the end of this article.

understand the sentence describing the phenomenon to be explained (not that phenomenon itself); by the explanans, the class of those sentences which are adduced to account for the phenomenon. As was noted before, the explanans falls into two subclasses; one of these contains certain sentences $C_1, C_2, \cdots, C_k$ which state specific antecedent conditions; the other is a set of sentences $L_1, L_2, \cdots L_r$ which represent general laws.

If a proposed explanation is to be sound, its constituents have to satisfy certain conditions of adequacy, which may be divided into logical and empirical conditions. For the following discussion, it will be sufficient to formulate these requirements in a slightly vague manner; in Part III, a more rigorous anlysis and a more precise restatement of these criteria will be presented.

I. *Logical conditions of adequacy.*

(R1) The explanandum must be a logical consequence of the explanans; in other words, the explanandum must be logically deducible from the information contained in the explanans, for otherwise, the explanans would not constitute adequate grounds for the explanandum.

(R2) The explanans must contain general laws, and these must actually be required for the derivation of the explanandum.—We shall not make it a necessary condition for a sound explanation, however, that the explanans must contain at least one statement which is not a law; for, to mention just one reason, we would surely want to consider as an explanation the derivation of the general regularities governing the motion of double stars from the laws of celestial mechanics, even though all the statements in the explanans are general laws.

(R3) The explanans must have empirical content; i.e., it must be capable, at least in principle, of test by experiment or observation.—This condition is implicit in (R1); for since the explanandum is assumed to describe some empirical phenomenon, it follows from (R1) that the explanans entails at least one consequence of empirical character, and this fact confers upon it testability and empirical content. But the point deserves special mention because, as will be seen in §4, certain arguments which have been offered as explanations in the natural and in the social sciences violate this requirement.

II. *Empirical condition of adequacy.*

(R4) The sentences constituting the explanans must be true.

That in a sound explanation, the statements constituting the explanans have to satisfy some condition of factual correctness is obvious. But it might seem more appropriate to stipulate that the explanans has to be highly confirmed by all the relevant evidence available rather than that it should be true. This stipulation however, leads to awkward consequences. Suppose that a certain phenomenon was explained at an earlier stage of science, by means of an explanans which was well supported by the evidence then at hand, but which had been highly disconfirmed by more recent empirical findings. In such a case, we would have to say that originally the explanatory account was a correct explanation, but that it ceased to be one later, when unfavorable evidence was discovered. This does not appear to accord with sound common usage, which directs us to say that on the basis of the limited initial evidence, the truth of the explanans, and thus the soundness of the explanation, had been quite probable, but that the ampler evidence now available made it highly probable that the explanans was not true, and hence that the account in question was not—and had never been—a correct explanation. (A similar point will be made and illustrated, with respect to the requirement of truth for laws, in the beginning of §6.)

Some of the characteristics of an explanation which have been indicated so far may be summarized in the following schema:

Logical deduction	$C_1, C_2, \cdots, C_k$	Statements of antecedent conditions	Explanans
	$L_1, L_2, \cdots, L_r$	General Laws	
→	E	Description of the empirical phenomenon to be explained	Explanandum

Let us note here that the same formal analysis, including the four necessary conditions, applies to scientific prediction as well as to explanation. The difference between the two is of a pragmatic character. If E is given, i.e. if we know that the phenomenon described by E has occurred, and a suitable set of statements $C_1, C_2, \cdots, C_k, L_1, L_2, \cdots, L_r$ is provided afterwards, we speak of an explanation of the phenomenon in question. If the latter statements are given and E is derived prior to the occurrence of the phenomenon it describes, we speak of a prediction. It may be said, therefore, that an explanation is not fully adequate unless its explanans, if taken account of in time, could have served as a basis for predicting the phenomenon under consideration.[2a]—Consequently, whatever will be said in this article concerning the logical characteristics of explanation or prediction will be applicable to either, even if only one of them should be mentioned.

It is this potential predictive force which gives scientific explanation its importance: only to the extent that we are able to explain empirical facts can we attain the major objective of scientific research, namely not merely to record the phenomena of our experience, but to learn from them, by basing upon them theoretical generalizations which enable us to anticipate new occurrences and to control, at least to some extent, the changes in our environment.

Many explanations which are customarily offered, especially in pre-scientific discourse, lack this predictive character, however. Thus, it may be explained

[2a] The logical similarity of explanation and prediction, and the fact that one is directed towards past occurrences, the other towards future ones, is well expressed in the terms "postdictability" and "predictability" used by Reichenbach in [Quantum Mechanics], p. 13.

that a car turned over on the road "because" one of its tires blew out while the car was travelling at high speed. Clearly, on the basis of just this information, the accident could not have been predicted, for the explanans provides no explicit general laws by means of which the prediction might be effected, nor does it state adequately the antecedent conditions which would be needed for the prediction.—The same point may be illustrated by reference to W. S. Jevons's view that every explanation consists in pointing out a resemblance between facts, and that in some cases this process may require no reference to laws at all and "may involve nothing more than a single identity, as when we explain the appearance of shooting stars by showing that they are identical with portions of a comet".[8] But clearly, this identity does not provide an explanation of the phenomenon of shooting stars unless we presuppose the laws governing the development of heat and light as the effect of friction. The observation of similarities has explanatory value only if it involves at least tacit reference to general laws.

In some cases, incomplete explanatory arguments of the kind here illustrated suppress parts of the explanans simply as "obvious"; in other cases, they seem to involve the assumption that while the missing parts are not obvious, the incomplete explanans could at least, with appropriate effort, be so supplemented as to make a strict derivation of the explanandum possible. This assumption may be justifiable in some cases, as when we say that a lump of sugar disappeared "because" it was put into hot tea, but it is surely not satisfied in many other cases. Thus, when certain peculiarities in the work of an artist are explained as outgrowths of a specific type of neurosis, this observation may contain significant clues, but in general it does not afford a sufficient basis for a potential prediction of those peculiarities. In cases of this kind, an incomplete explanation may at best be considered as indicating some positive correlation between the antecedent conditions adduced and the type of phenomenon to be explained, and as pointing out a direction in which further research might be carried on in order to complete the explanatory account.

The type of explanation which has been considered here so far is often referred to as causal explanation. If E describes a particular event, then the antecedent circumstances described in the sentences $C_1, C_2, \cdots, C_k$ may be said jointly to "cause" that event, in the sense that there are certain empirical regularities, expressed by the laws $L_1, L_2, \cdots, L_r$, which imply that whenever conditions of the kind indicated by $C_1, C_2, \cdots, C_k$ occur, an event of the kind described in E will take place. Statements such as $L_1, L_2, \cdots, L_r$, which assert general and unexceptional connections between specified characteristics of events, are customarily called causal, or deterministic, laws. They are to be distinguished from the so-called statistical laws which assert that in the long run, an explicitly stated percentage of all cases satisfying a given set of conditions are accompanied by an event of a certain specified kind. Certain cases of scientific explanation involve "subsumption" of the explanandum under a set of laws of which at least some are statistical in character. Analysis of the peculiar logical structure

[8] [Principles], p. 533.

of that type of subsumption involves difficult special problems. The present essay will be restricted to an examination of the causal type of explanation, which has retained its significance in large segments of contemporary science, and even in some areas where a more adequate account calls for reference to statistical laws.[4]

§4. *Explanation in the non-physical sciences. Motivational and teleological approaches.*

Our characterization of scientific explanation is so far based on a study of cases taken from the physical sciences. But the general principles thus obtained apply also outside this area.[5] Thus, various types of behavior in laboratory animals and in human subjects are explained in psychology by subsumption under laws or even general theories of learning or conditioning; and while frequently, the regularities invoked cannot be stated with the same generality and precision as in physics or chemistry, it is clear, at least, that the general character of those explanations conforms to our earlier characterization.

Let us now consider an illustration involving sociological and economic factors. In the fall of 1946, there occurred at the cotton exchanges of the United States a price drop which was so severe that the exchanges in New York, New Orleans, and Chicago had to suspend their activities temporarily. In an attempt to explain this occurrence, newspapers traced it back to a large-scale speculator in New Orleans who had feared his holdings were too large and had therefore begun to liquidate his stocks; smaller speculators had then followed his example

[4] The account given above of the general characteristics of explanation and prediction in science is by no means novel; it merely summarizes and states explicitly some fundamental points which have been recognized by many scientists and methodologists.

Thus, e.g., Mill says: "An individual fact is said to be explained by pointing out its cause, that is, by stating the law or laws of causation of which its production is an instance", and "a law of uniformity in nature is said to be explained when another law or laws are pointed out, of which that law itself is but a case, and from which it could be deduced." ([Logic], Book III, Chapter XII, section 1). Similarly, Jevons, whose general characterization of explanation was critically discussed above, stresses that "the most important process of explanation consists in showing that an observed fact is one case of a general law or tendency." ([Principles], p. 533). Ducasse states the same point as follows: "Explanation essentially consists in the offering of a hypothesis of fact, standing to the fact to be explained as case of antecedent to case of consequent of some already known law of connection." ([Explanation], pp. 150–51). A lucid analysis of the fundamental structure of explanation and prediction was given by Popper in [Forschung], section 12, and, in an improved version, in his work [Society], especially in Chapter 25 and in note 7 referring to that chapter.—For a recent characterization of explanation as subsumption under general theories, cf., for example, Hull's concise discussion in [Principles], chapter I. A clear elementary examination of certain aspects of explanation is given in Hospers, [Explanation], and a concise survey of many of the essentials of scientific explanation which are considered in the first two parts of the present study may be found in Feigl, [Operationism], pp. 284 ff.

[5] On the subject of explanation in the social sciences, especially in history, cf. also the following publications, which may serve to supplement and amplify the brief discussion to be presented here: Hempel, [Laws]; Popper, [Society]; White, [Explanation]; and the articles *Cause* and *Understanding* in Beard and Hook, [Terminology].

in a panic and had thus touched off the critical decline. Without attempting to assess the merits of the argument, let us note that the explanation here suggested again involves statements about antecedent conditions and the assumption of general regularities. The former include the facts that the first speculator had large stocks of cotton, that there were smaller speculators with considerable holdings, that there existed the institution of the cotton exchanges with their specific mode of operation, etc. The general regularities referred to are—as often in semi-popular explanations—not explicitly mentioned; but there is obviously implied some form of the law of supply and demand to account for the drop in cotton prices in terms of the greatly increased supply under conditions of practically unchanged demand; besides, reliance is necessary on certain regularities in the behavior of individuals who are trying to preserve or improve their economic position. Such laws cannot be formulated at present with satisfactory precision and generality, and therefore, the suggested explanation is surely incomplete, but its intention is unmistakably to account for the phenomenon by integrating it into a general pattern of economic and socio-psychological regularities.

We turn to an explanatory argument taken from the field of linguistics.[6] In Northern France, there exist a large variety of words synonymous with the English "bee," whereas in Southern France, essentially only one such word is in existence. For this discrepancy, the explanation has been suggested that in the Latin epoch, the South of France used the word "apicula", the North the word "apis". The latter, because of a process of phonologic decay in Northern France, became the monosyllabic word "é"; and monosyllables tend to be eliminated, especially if they contain few consonantic elements, for they are apt to give rise to misunderstandings. Thus, to avoid confusion, other words were selected. But "apicula", which was reduced to "abelho", remained clear enough and was retained, and finally it even entered into the standard language, in the form "abbeille". While the explanation here described is incomplete in the sense characterized in the previous section, it clearly exhibits reference to specific antecedent conditions as well as to general laws.[7]

While illustrations of this kind tend to support the view that explanation in biology, psychology, and the social sciences has the same structure as in the physical sciences, the opinion is rather widely held that in many instances, the causal type of explanation is essentially inadequate in fields other than physics and chemistry, and especially in the study of purposive behavior. Let us ex-

[6] The illustration is taken from Bonfante, [Semantics], section 3.

[7] While in each of the last two illustrations, certain regularities are unquestionably relied upon in the explanatory argument, it is not possible to argue convincingly that the intended laws, which at present cannot all be stated explicitly, are of a causal rather than a statistical character. It is quite possible that most or all of the regularities which will be discovered as sociology develops will be of a statistical type. Cf., on this point, the suggestive observations by Zilsel in [Empiricism] section 8, and [Laws]. This issue does not affect, however, the main point we wish to make here, namely that in the social no less than in the physical sciences, subsumption under general regularities is indispensable for the explanation and the theoretical understanding of any phenomenon.

amine briefly some of the reasons which have been adduced in support of this view.

One of the most familiar among them is the idea that events involving the activities of humans singly or in groups have a peculiar uniqueness and irrepeatability which makes them inaccessible to causal explanation because the latter, which its reliance upon uniformities, presupposes repeatability of the phenomena under consideration. This argument which, incidentally, has also been used in support of the contention that the experimental method is inapplicable in psychology and the social sciences, involves a misunderstanding of the logical character of causal explanation. Every individual event, in the physical sciences no less than in psychology or the social sciences, is unique in the sense that it, with all its peculiar characteristics, does not repeat itself. Nevertheless, individual events may conform to, and thus be explainable by means of, general laws of the causal type. For all that a causal law asserts is that any event of a specified kind, i.e. any event having certain specified characteristics, is accompanied by another event which in turn has certain specified characteristics; for example, that in any event involving friction, heat is developed. And all that is needed for the testability and applicability of such laws is the recurrence of events with the antecedent characteristics, i.e. the repetition of those characteristics, but not of their individual instances. Thus, the argument is inconclusive. It gives occasion, however, to emphasize an important point concerning our earlier analysis: When we spoke of the explanation of a single event, the term "event" referred to the occurrence of some more or less complex characteristic in a specific spatio-temporal location or in a certain individual object, and not to *all* the characteristics of that object, or to all that goes on in that space-time region.

A second argument that should be mentioned here[8] contends that the establishment of scientific generalizations—and thus of explanatory principles—for human behavior is impossible because the reactions of an individual in a given situation depend not only upon that situation, but also upon the previous history of the individual.—But surely, there is no *a priori* reason why generalizations should not be attainable which take into account this dependence of behavior on the past history of the agent. That indeed the given argument "proves" too much, and is therefore a *non sequitur*, is made evident by the existence of certain physical phenomena, such as magnetic hysteresis and elastic fatigue, in which the magnitude of a specific physical effect depends upon the past history of the system involved, and for which nevertheless certain general regularities have been established.

A third argument insists that the explanation of any phenomenon involving purposive behavior calls for reference to motivations and thus for teleological rather than causal analysis. Thus, for example, a fuller statement of the suggested explanation for the break in the cotton prices would have to indicate the large-scale speculator's motivations as one of the factors determining the event

[8] Cf., for example, F. H. Knight's presentation of this argument in [Limitations], pp. 251–52.

in question. Thus, we have to refer to goals sought, and this, so the argument runs, introduces a type of explanation alien to the physical sciences. Unquestionably, many of the—frequently incomplete—explanations which are offered for human actions involve reference to goals and motives; but does this make them essentially different from the causal explanations of physics and chemistry? One difference which suggests itself lies in the circumstance that in motivated behavior, the future appears to affect the present in a manner which is not found in the causal explanations of the physical sciences. But clearly, when the action of a person is motivated, say, by the desire to reach a certain objective, then it is not the as yet unrealized future event of attaining that goal which can be said to determine his present behavior, for indeed the goal may never be actually reached; rather—to put it in crude terms—it is (a) his desire, present before the action, to attain that particular objective, and (b) his belief, likewise present before the action, that such and such a course of action is most likely to have the desired effect. The determining motives and beliefs, therefore, have to be classified among the antecedent conditions of a motivational explanation, and there is no formal difference on this account between motivational and causal explanation.

Neither does the fact that motives are not accessible to direct observation by an outside observer constitute an essential difference between the two kinds of explanation; for also the determining factors adduced in physical explanations are very frequently inaccessible to direct observation. This is the case, for instance, when opposite electric charges are adduced in explanation of the mutual attraction of two metal spheres. The presence of those charges, while eluding all direct observation, can be ascertained by various kinds of indirect test, and that is sufficient to guarantee the empirical character of the explanatory statement. Similarly, the presence of certain motivations may be ascertainable only by indirect methods, which may include reference to linguistic utterances of the subject in question, slips of the pen or of the tongue, etc.; but as long as these methods are "operationally determined" with reasonable clarity and precision, there is no essential difference in this respect between motivational explanation and causal explanation in physics.

A potential danger of explanation by motives lies in the fact that the method lends itself to the facile construction of ex-post-facto accounts without predictive force. It is a widespread tendency to "explain" an action by ascribing it to motives conjectured only after the action has taken place. While this procedure is not in itself objectionable, its soundness requires that (1) the motivational assumptions in question be capable of test, and (2) that suitable general laws be available to lend explanatory power to the assumed motives. Disregard of these requirements frequently deprives alleged motivational explanations of their cognitive significance.

The explanation of an action in terms of the motives of the agent is sometimes considered as a special kind of teleological explanation. As was pointed out above, motivational explanation, if adequately formulated, conforms to the conditions for causal explanation, so that the term "teleological" is a misnomer if it is meant to imply either a non-causal character of the explanation or a peculiar determination of the present by the future. If this is borne in mind, however, the term "teleological" may be viewed, in this context, as referring to causal explanations in which some of the antecedent conditions are motives of the agent whose actions are to be explained.[9]

Teleological explanations of this kind have to be distinguished from a much more sweeping type, which has been claimed by certain schools of thought to be indispensable especially in biology. It consists in explaining characteristics of an organism by reference to certain ends or purposes which the characteristics are said to serve. In contradistinction to the cases examined before, the ends are not assumed here to be consciously or subconsciously pursued by the organism in question. Thus, for the phenomenon of mimicry, the explanation is sometimes offered that it serves the purpose of protecting the animals endowed with it from detection by its pursuers and thus tends to preserve the species. —Before teleological hypotheses of this kind can be appraised as to their potential explanatory power, their meaning has to be clarified. If they are intended somehow to express the idea that the purposes they refer to are inherent in the design of the universe, then clearly they are not capable of empirical test and thus violate the requirement (R3) stated in §3. In certain cases, however, assertions about the purposes of biological characteristics may be translatable into statements in non-teleological terminology which assert that those characteristics function in a specific manner which is essential to keeping the organism alive or to preserving the species.[10] An attempt to state precisely what is meant by this latter assertion—or by the similar one that without those characteristics, and other things being equal, the organism or the species would not survive—encounters considerable difficulties. But these need not be discussed here. For even if we assume that biological statements in teleological form can be adequately translated into descriptive statements about the life-preserving function of certain biological characteristics, it is clear that (1) the use of the concept of purpose is not essential in these contexts, since the term "purpose" can be completely eliminated from the statements in question, and (2) teleological assumptions, while now endowed with empirical content, cannot serve as explanatory principles in the customary contexts. Thus, e.g., the fact that a

[9] For a detailed logical analysis of the character and the function of the motivation concept in psychological theory, see Koch, [Motivation].—A stimulating discussion of teleological behavior from the standpoint of contemporary physics and biology is contained in the article [Teleology] by Rosenblueth, Wiener and Bigelow. The authors propose an interpretation of the concept of purpose which is free from metaphysical connotations, and they stress the importance of the concept thus obtained for a behavioristic analysis of machines and living organisms. While our formulations above intentionally use the crude terminology frequently applied in philosophical arguments concerning the applicability of causal explanation to purposive behavior, the analysis presented in the article referred to is couched in behavioristic terms and avoids reference to "motives" and the like.

[10] An analysis of teleological statements in biology along these lines may be found in Woodger, [Principles], especially pp. 432 ff; essentially the same interpretation is advocated by Kaufmann in [Methodology], chapter 8.

given species of butterflies displays a particular kind of coloring cannot be inferred from—and therefore cannot be explained by means of—the statement that this type of coloring has the effect of protecting the butterflies from detection by pursuing birds, nor can the presence of red corpuscles in the human blood be inferred from the statement that those corpuscles have a specific function in assimilating oxygen and that this function is essential for the maintenance of life.

One of the reasons for the perseverance of teleological considerations in biology probably lies in the fruitfulness of the teleological approach as a heuristic device: Biological research which was psychologically motivated by a teleological orientation, by an interest in purposes in nature, has frequently led to important results which can be stated in non-teleological terminology and which increase our scientific knowledge of the causal connections between biological phenomena.

Another aspect that lends appeal to teleological considerations is their anthropomorphic character. A teleological explanation tends to make us feel that we really "understand" the phenomenon in question, because it is accounted for in terms of purposes, with which we are familiar from our own experience of purposive behavior. But it is important to distinguish here understanding in the psychological sense of a feeling of empathic familiarity from understanding in the theoretical, or cognitive, sense of exhibiting the phenomenon to be explained as a special case of some general regularity. The frequent insistence that explanation means the reduction of something unfamiliar to ideas or experiences already familiar to us is indeed misleading. For while some scientific explanations do have this psychological effect, it is by no means universal: The free fall of a physical body may well be said to be a more familiar phenomenon than the law of gravitation, by means of which it can be explained; and surely the basic ideas of the theory of relativity will appear to many to be far less familiar than the phenomena for which the theory accounts.

"Familiarity" of the explicans is not only not necessary for a sound explanation—as we have just tried to show—, but it is not sufficient either. This is shown by the many cases in which a proposed explicans sounds suggestively familiar, but upon closer inspection proves to be a mere metaphor, or an account lacking testability, or a set of statements which includes no general laws and therefore lacks explanatory power. A case in point is the neovitalistic attempt to explain biological phenomena by reference to an entelechy or vital force. The crucial point here is not—as it is sometimes made out to be—that entelechies cannot be seen or otherwise directly observed; for that is true also of gravitational fields, and yet, reference to such fields is essential in the explanation of various physical phenomena. The decisive difference between the two cases is that the physical explanation provides (1) methods of testing, albeit indirectly, assertions about gravitational fields, and (2) general laws concerning the strength of gravitational fields, and the behavior of objects moving in them. Explanations by entelechies satisfy the analogue of neither of these two conditions. Failure to satisfy the first condition represents a violation of (R3); it renders all statements about entelechies inaccessible to empirical test and thus devoid of empirical meaning. Failure to comply with the second condition involves a violation of (R2). It deprives the concept of entelechy of all explanatory import; for explanatory power never resides in a concept, but always in the general laws in which it functions. Therefore, notwithstanding the flavor of familiarity of the metaphor it invokes, the neovitalistic approach cannot provide theoretical understanding.

The preceding observations about familiarity and understanding can be applied, in a similar manner, to the view held by some scholars that the explanation, or the understanding, of human actions requires an empathic understanding of the personalities of the agents[11]. This understanding of another person in terms of one's own psychological functioning may prove a useful heuristic device in the search for general psychological principles which might provide a theoretical explanation; but the existence of empathy on the part of the scientist is neither a necessary nor a sufficient condition for the explanation, or the scientific understanding, of any human action. It is not necessary, for the behavior of psychotics or of people belonging to a culture very different from that of the scientist may sometimes be explainable and predictable in terms of general principles even though the scientist who establishes or applies those principles may not be able to understand his subjects empathically. And empathy is not sufficient to guarantee a sound explanation, for a strong feeling of empathy may exist even in cases where we completely misjudge a given personality. Moreover, as the late Dr. Zilsel has pointed out, empathy leads with ease to incompatible results; thus, when the population of a town has long been subjected to heavy bombing attacks, we can understand, in the empathic sense, that its morale should have broken down completely, but we can understand with the same ease also that it should have developed a defiant spirit of resistance. Arguments of this kind often appear quite convincing; but they are of an *ex post facto* character and lack cognitive significance unless they are supplemented by testable explanatory principles in the form of laws or theories.

Familiarity of the explanans, therefore, no matter whether it is achieved through the use of teleological terminology, through neovitalistic metaphors, or through other means, is no indication of the cognitive import and the predictive force of a proposed explanation. Besides, the extent to which an idea will be considered as familiar varies from person to person and from time to time, and a psychological factor of this kind certainly cannot serve as a standard in assessing the worth of a proposed explanation. The decisive requirement for every sound explanation remains that it subsume the explanandum under general laws.

PART II. ON THE IDEA OF EMERGENCE

§5. *Levels of Explanation. Analysis of Emergence.*

As has been shown above, a phenomenon may often be explained by sets of laws of different degrees of generality. The changing positions of a planet, for example, may be explained by subsumption under Kepler's laws, or by deriva-

[11] For a more detailed discussion of this view on the basis of the general principles outlined above, cf. Zilsel, [Empiricism], sections 7 and 8, and Hempel, [Laws], section 6.

tion from the far more comprehensive general law of gravitation in combination with the laws of motion, or finally by deduction from the general theory of relativity, which explains—and slightly modifies—the preceding set of laws. Similarly, the expansion of a gas with rising temperature at constant pressure may be explained by means of the Gas Law or by the more comprehensive kinetic theory of heat. The latter explains the Gas Law, and thus indirectly the phenomenon just mentioned, by means of (1) certain assumptions concerning the micro-behavior of gases (more specifically, the distributions of locations and speeds of the gas molecules) and (2) certain macro-micro principles, which connect such macro-characteristics of a gas as its temperature, pressure and volume with the micro-characteristics just mentioned.

In the sense of these illustrations, a distinction is frequently made between various levels of explanation[12]. Subsumption of a phenomenon under a general law directly connecting observable characteristics represents the first level; higher levels require the use of more or less abstract theoretical constructs which function in the context of some comprehensive theory. As the preceding illustrations show, the concept of higher-level explanation covers procedures of rather different character; one of the most important among them consists in explaining a class of phenomena by means of a theory concerning their micro-structure. The kinetic theory of heat, the atomic theory of matter, the electromagnetic as well as the quantum theory of light, and the gene theory of heredity are examples of this method. It is often felt that only the discovery of a micro-theory affords real scientific understanding of any type of phenomenon, because only it gives us insight into the inner mechanism of the phenomenon, so to speak. Consequently, classes of events for which no micro-theory was available have frequently been viewed as not actually understood; and concern with the theoretical status of phenomena which are unexplained in this sense may be considered as a theoretical root of the doctrine of emergence.

Generally speaking, the concept of emergence has been used to characterize certain phenomena as "novel", and this not merely in the psychological sense of being unexpected[13], but in the theoretical sense of being unexplainable, or unpredictable, on the basis of information concerning the spatial parts or other constituents of the systems in which the phenomena occur, and which in this context are often referred to as wholes. Thus, e.g., such characteristics of water as its transparence and liquidity at room temperature and atmospheric pressure, or its ability to quench thirst have been considered as emergent on the ground that they could not possibly have been predicted from a knowledge of the properties of its chemical constituents, hydrogen and oxygen. The weight of the compound, on the contrary, has been said not to be emergent because it is a mere "resultant" of its components and could have been predicted by simple addition even before the compound had been formed. The conceptions of ex-

[12] For a lucid brief exposition of this idea, see Feigl, [Operationism], pp. 284–288.

[13] Concerning the concept of novelty in its logical and psychological meanings, see also Stace, [Novelty].

planation and prediction which underly this idea of emergence call for various critical observations, and for corresponding changes in the concept of emergence.

(1) First, the question whether a given characteristic of a "whole", w, is emergent or not cannot be significantly raised until it has been stated what is to be understood by the parts or constituents of w. The volume of a brick wall, for example, may be inferable by addition from the volumes of its parts if the latter are understood to be the component bricks, but it is not so inferable from the volumes of the molecular components of the wall. Before we can significantly ask whether a characteristic W of an object w is emergent, we shall therefore have to state the intended meaning of the term "part of". This can be done by defining a specific relation Pt and stipulating that those and only those objects which stand in Pt to w count as parts or constituents of w. 'Pt' might be defined as meaning "constituent brick of" (with respect to buildings), or "molecule contained in" (for any physical object), or "chemical element contained in" (with respect to chemical compounds, or with respect to any material object), or "cell of" (with respect to organisms), etc. The term "whole" will be used here without any of its various connotations, merely as referring to any object w to which others stand in the specified relation Pt. In order to emphasize the dependence of the concept of part upon the definition of the relation Pt in each case, we shall sometimes speak of Pt-parts, to refer to parts as determined by the particular relation Pt under consideration.

(2) We turn to a second point of criticism. If a characteristic of a whole is to be qualified as emergent only if its occurrence cannot be inferred from a knowledge of all the properties of its parts, then, as Grelling has pointed out, no whole can have any emergent characteristics. Thus, to illustrate by reference to our earlier example, the properties of hydrogen include that of forming, if suitably combined with oxygen, a compound which is liquid, transparent, etc. Hence the liquidity, transparence, etc. of water can be inferred from certain properties of its chemical constituents. If the concept of emergence is not to be vacuous, therefore, it will be necessary to specify in every case a class G of attributes and to call a characteristic W of an object w emergent relatively to G and Pt if the occurrence of W in w cannot be inferred from a complete characterization of all the Pt-parts with respect to the attributes contained in G, i.e. from a statement which indicates, for every attribute in G, to which of the parts of w it applies. —Evidently, the occurrence of a characteristic may be emergent with respect to one class of attributes and not emergent with respect to another. The classes of attributes which the emergentists have in mind, and which are usually not explicitly indicated, will have to be construed as non-trivial, i.e. as not logically entailing the property of each constituent of forming, together with the other constituents, a whole with the characteristics under investigations.—Some fairly simple cases of emergence in the sense so far specified arise when the class G is restricted to certain simple properties of the parts, to the exclusion of spatial or other relations among them. Thus, the electromotive force of a system of several electric batteries cannot be inferred from the electromotive forces of its

constituents alone without a description, in terms of relational concepts, of the way in which the batteries are connected with each other.[14]

(3) Finally, the predictability of a given characteristic of an object on the basis of specified information concerning its parts will obviously depend on what general laws or theories are available.[15] Thus, the flow of an electric current in a wire connecting a piece of copper and a piece of zinc which are partly immersed in sulfuric acid is unexplainable, on the basis of information concerning any non-trivial set of attributes of copper, zinc and sulfuric acid, and the particular structure of the system under consideration, unless the theory available contains certain general laws concerning the functioning of batteries, or even more comprehensive principles of physical chemistry. If the theory includes such laws, on the other hand, then the occurrence of the current is predictable. Another illustration, which at the same time provides a good example for the point made under (2) above, is afforded by the optical activity of certain substances. The optical activity of sarco-lactic acid, for example, i.e. the fact that in solution it rotates the plane of polarization of plane-polarized light, cannot be predicted on the basis of the chemical characteristics of its constituent elements; rather, certain facts about the relations of the atoms constituting a molecule of sarco-lactic acid have to be known. The essential point is that the molecule in question contains an asymmetric carbon atom, i.e. one that holds four different atoms or groups, and if this piece of relational information is provided, the optical activity of the solution can be predicted provided that furthermore the theory available for the purpose embodies the law that the presence of one asymmetric carbon atom in a molecule implies optical activity of the solution; if the theory does not include this micro-macro law, then the phenomenon is emergent with respect to that theory.

An argument is sometimes advanced to the effect that phenomena such as the

[14] This observation connects the present discussion with a basic issue in Gestalt theory. Thus, e.g., the insistence that "a whole is more than the sum of its parts" may be construed as referring to characteristics of wholes whose prediction requires knowledge of certain structural relations among the parts. For a further examination of this point, see Grelling and Oppenheim, [Gestaltbegriff] and [Functional Whole].

[15] Logical analyses of emergence which make reference to the theories available have been propounded by Grelling and recently, in a very explicit form, by Henle in [Emergence]. In effect, Henle's definition characterizes a phenomenon as emergent if it cannot be predicted, by means of the theories accepted at the time, on the basis of the data available before its occurrence. In this interpretation of emergence, no reference is made to characteristics of parts or constitutents. Henle's concept of predictability differs from the one implicit in our discussion (and made explicit in Part III of this article) in that it implies derivability from the "simplest" hypothesis which can be formed on the basis of the data and theories available at the time. A number of suggestive observations on the idea of emergence and on Henle's analysis of it are contained in Bergmann's article [Emergence].—The idea that the concept of emergence, at least in some of its applications, is meant to refer to unpredictability by means of "simple" laws was advanced also by Grelling in the correspondence mentioned in note (1). Reliance on the motion of simplicity of hypotheses, however, involves considerable difficulties; in fact, no satisfactory definition of that concept is available at present.

flow of the current, or the optical activity, in our last examples, are absolutely emergent at least in the sense that they could not possibly have been predicted before they had been observed for the first time; in other words, that the laws requisite for their prediction could not have been arrived at on the basis of information available before their first observed occurrence.[16] This view is untenable, however. On the strength of data available at a given time, science often establishes generalizations by means of which it can forecast the occurrence of events the like of which have never before been encountered. Thus, generalizations based upon periodicities exhibited by the characteristics of chemical elements then known, enabled Mendeleeff in 1871 to predict the existence of a certain new element and to state correctly various properties of that element as well as of several of its compounds; the element in question, germanium, was not discovered until 1886.—A more recent illustration of the same point is provided by the development of the atomic bomb and the prediction, based on theoretical principles established prior to the event, of its explosion under specified conditions, and of its devastating release of energy.

As Grelling has stressed, the observation that the predictability of the occurrence of any characteristic depends upon the theoretical knowledge available, applies even to those cases in which, in the language of some emergentists, the characteristic of the whole is a mere resultant of the corresponding characteristics of the parts and can be obtained from the latter by addition. Thus, even the weight of a water molecule cannot be derived from the weights of its atomic constituents without the aid of a law which expresses the former as some specific mathematical function of the latter. That this function should be the sum is by no means self-evident; it is an empirical generalization, and at that not a strictly correct one, as relativistic physics has shown.

Failure to realize that the question of the predictability of a phenomenon cannot be significantly raised unless the theories available for the prediction have been specified has encouraged the misconception that certain phenomena have a mysterious quality of absolute unexplainability, and that their emergent status has to be accepted with "natural piety", as F. L. Morgan put it. The observations presented in the preceding discussion strip the idea of emergence of these unfounded connotations: emergence of a characteristic is not an ontological trait inherent in some phenomena; rather it is indicative of the scope of our knowl

[16] C. D. Broad, who in chapter 2 of his book, [Mind], gives a clear presentation and critical discussion of the essentials of emergentism, emphasizes the importance of "laws of composition" in predicting the characteristics of a whole on the basis of those of its parts. (cf. [Mind], pp. 61ff.); but he subscribes to the view characterized above and illustrates it specifically by the assertion that "if we want to know the chemical (and many of the physical) properties of a chemical compound, such as silver-chloride, it is absolutely necessary to study samples of *that particular compound*. . . . The essential point is that it would also be useless to study chemical compounds in general and to compare their properties with those of their elements in the hope of discovering a *general* law of composition by which the properties of *any* chemical compound could be foretold when the properties of its separate elements were known." (Ibid., p. 64)—That an achievement of precisely this sort has been possible on the basis of the periodic system of the elements is pointed out above.

edge at a given time; thus it has no absolute, but a relative character; and what is emergent with respect to the theories available today may lose its emergent status tomorrow.

The preceding considerations suggest the following redefinition of emergence: The occurrence of a characteristic W in an object w is emergent relatively to a theory T, a part relation Pt, and a class G of attributes if that occurrence cannot be deduced by means of T from a characterization of the Pt-parts of w with respect to all the attributes in G.

This formulation explicates the meaning of emergence with respect to *events* of a certain kind, namely the occurrence of some characteristic W in an object w. Frequently, emergence is attributed to *characteristics* rather than to events; this use of the concept of emergence may be interpreted as follows: A characteristic W is emergent relatively to T, Pt, and G if its occurrence in *any* object is emergent in the sense just indicated.

As far as its cognitive content is concerned, the emergentist assertion that the phenomena of life are emergent may now be construed, roughly, as an elliptic formulation of the following statement: Certain specifiable biological phenomena cannot be explained, by means of contemporary physico-chemical theories, on the basis of data concerning the physical and chemical characteristics of the atomic and molecular constituents of organisms. Similarly, the so-called emergent status of mind reduces to the assertion that present-day physical, chemical and biological theories do not suffice to explain all psychological phenomena on the basis of data concerning the physical, chemical, and biological characteristics of the cells or of the molecules or atoms constituting the organisms in question. But in this interpretation, the emergent character of biological and psychological phenomena becomes trivial; for the description of various biological phenomena requires terms which are not contained in the vocabulary of present day physics and chemistry; hence we cannot expect that all specifically biological phenomena are explainable, i.e. deductively inferable, by means of present day physico-chemical theories on the basis of initial conditions which themselves are described in exclusively physico-chemical terms. In order to obtain a less trivial interpretation of the assertion that the phenomena of life are emergent, we have therefore to include in the explanatory theory all those laws known at present which connect the physico-chemical with the biological "level", i.e., which contain, on the one hand, certain physical and chemical terms, including those required for the description of molecular structures, and on the other hand, certain concepts of biology. An analogous observation applies to the case of psychology. If the assertion that life and mind have an emergent status is interpreted in this sense, then its import can be summarized approximately by the statement that no explanation, in terms of micro-structure theories, is available at present for large classes of phenomena studied in biology and psychology.[17]

[17] The following passage from Tolman, [Behavior], may serve to support this interpretation: ". . . 'behavior-acts', though no doubt in complete one-to-one correspondence with the underlying molecular facts of physics and physiology, have, as 'molar' wholes, certain emergent properties of their own. . . . Further, these molar properties of behavior-acts

Assertions of this type, then, appear to represent the rational core of the doctrine of emergence. In its revised form, the idea of emergence no longer carries with it the connotation of absolute unpredictability—a notion which is objectionable not only because it involves and perpetuates certain logical misunderstandings, but also because, not unlike the ideas of neo-vitalism, it encourages an attitude of resignation which is stifling for scientific research. No doubt it is this characteristic, together with its theoretical sterility, which accounts for the rejection, by the majority of contemporary scientists, of the classical absolutistic doctrine of emergence.[18]

PART III. LOGICAL ANALYSIS OF LAW AND EXPLANATION

§6. *Problems of the concept of general law.*

From our general survey of the characteristics of scientific explanation, we now turn to a closer examination of its logical structure. The explanation of a phenomenon, we noted, consists in its subsumption under laws or under a theory. But what is a law, what is a theory? While the meaning of these concepts seems intuitively clear, an attempt to construct adequate explicit definitions for them encounters considerable difficulties. In the present section, some basic problems of the concept of law will be described and analyzed; in the next section, we intend to propose, on the basis of the suggestions thus obtained, definitions of law and of explanation for a formalized model language of a simple logical structure.

The concept of law will be construed here so as to apply to true statements only. The apparently plausible alternative procedure of requiring high confirmation rather than truth of a law seems to be inadequate: It would lead to a relativized concept of law, which would be expressed by the phrase "sentence S is a law relatively to the evidence E". This does not seem to accord with the meaning customarily assigned to the concept of law in science and in methodological inquiry. Thus, for example, we would not say that Bode's general formula for the distance of the planets from the sun was a law relatively to the astronomical evidence available in the 1770s, when Bode propounded it, and that it ceased to be a law after the discovery of Neptune and the determination of its distance from the sun; rather, we would say that the limited original evidence had given a high probability to the assumption that the formula was a law, whereas more recent additional information reduced that probability so much as to make it practically certain that Bode's formula is not generally true, and hence not a law.[18a]

---

cannot in the present state of our knowledge, i.e., prior to the working-out of many empirical correlations between behavior and its physiological correlates, be known even inferentially from a mere knowledge of the underlying, molecular, facts of physics and physiology." (l. c., pp. 7–8).—In a similar manner, Hull uses the distinction between molar and molecular theories and points out that theories of the latter type are not at present available in psychology. Cf. [Principles], pp. 19ff.; [Variables], p. 275.

[18] This attitude of the scientist is voiced, for example, by Hull in [Principles], pp. 24–28.

[18a] The requirement of truth for laws has the consequence that a given empirical statement S can never be definitely known to be a law; for the sentence affirming the truth of S

Apart from being true, a law will have to satisfy a number of additional conditions. These can be studied independently of the factual requirement of truth, for they refer, as it were, to all logically possible laws, no matter whether factually true or false. Adopting a convenient term proposed by Goodman[19], we will say that a sentence is lawlike if it has all the characteristics of a general law, with the possible exception of truth. Hence, every law is a lawlike sentence, but not conversely.

Our problem of analyzing the concept of law thus reduces to that of explicating the meaning of "lawlike sentence". We shall construe the class of lawlike sentences as including analytic general statements, such as "A rose is a rose", as well as the lawlike sentences of empirical science, which have empirical content.[20] It will not be necessary to require that each lawlike sentence permissible in explanatory contexts be of the second kind; rather, our definition of explanation will be so constructed as to guarantee the factual character of the totality of the laws—though not of every single one of them—which function in an explanation of an empirical fact.

What are the characteristics of lawlike sentences? First of all, lawlike sentences are statements of universal form, such as "All robins' eggs are greenish-blue", "All metals are conductors of electricity", "At constant pressure, any gas expands with increasing temperature". As these examples illustrate, a lawlike sentence usually is not only of universal, but also of conditional form; it makes an assertion to the effect that universally, if a certain set of conditions, C, is realized, then another specified set of conditions, E, is realized as well. The standard form for the symbolic expression of a lawlike sentence is therefore the universal conditional. However, since any conditional statement can be transformed into a non-conditional one, conditional form will not be considered as essential for a lawlike sentence, while universal character will be held indispensable.

But the requirement of universal form is not sufficient to characterize lawlike sentences. Suppose, for example, that a certain basket, b, contains at a certain time t a number of red apples and nothing else.[21] Then the statement

(S_1) Every apple in basket b at time t is red

is both true and of universal form. Yet the sentence does not qualify as a law; we would refuse, for example, to explain by subsumption under it the fact

---

is logically equivalent with S and is therefore capable only of acquiring a more or less high probability, or degree of confirmation, relatively to the experimental evidence available at any given time. On this point, cf. Carnap, [Remarks].—For an excellent non-technical exposition of the semantical concept of truth, which is here applied, the reader is referred to Tarski, [Truth].

[19] [Counterfactuals]. p. 125.

[20] This procedure was suggested by Goodman's approach in [Counterfactuals].—Reichenbach, in a detailed examination of the concept of law, similarly construes his concept of nomological statement as including both analytic and synthetic sentences; cf. [Logic], chapter VIII.

[21] The difficulty illustrated by this example was stated concisely by Langford ([Review]), who referred to it as the problem of distinguishing between universals of fact and causal universals. For further discussion and illustration of this point, see also Chisholm [Conditional], especially pp. 301f.—A systematic analysis of the problem was given by Goodman

that a particular apple chosen at random from the basket is red. What distinguishes S_1 from a lawlike sentence? Two points suggest themselves, which will be considered in turn, namely, finite scope, and reference to a specified object.

First, the sentence S_1 makes, in effect, an assertion about a finite number of objects only, and this seems irreconcilable with the claim to universality which is commonly associated with the notion of law.[22] But are not Kepler's laws considered as lawlike although they refer to a finite set of planets only? And might we not even be willing to consider as lawlike a sentence such as the following?

(S_2) All the sixteen ice cubes in the freezing tray of this refrigerator have a temperature of less than 10 degrees centigrade.

This point might well be granted; but there is an essential difference between S_1 on the one hand and Kepler's laws as well as S_2 on the other: The latter, while finite in scope, are known to be consequences of more comprehensive laws whose scope is not limited, while for S_1 this is not the case.

Adopting a procedure recently suggested by Reichenbach[23], we will therefore distinguish between fundamental and derivative laws. A statement will be called a derivative law if it is of universal character and follows from some fundamental laws. The concept of fundamental law requires further clarification; so far, we may say that fundamental laws, and similarly fundamental lawlike sentences, should satisfy a certain condition of non-limitation of scope.

It would be excessive, however, to deny the status of fundamental lawlike sentence to all statements which, in effect, make an assertion about a finite class of objects only, for that would rule out also a sentence such as "All robins' eggs are greenish-blue", since presumably the class of all robins' eggs—past, present, and future—is finite. But again, there is an essential difference between this sentence and, say, S_1. It requires empirical knowledge to establish the finiteness of the class of robins' eggs, whereas, when the sentence S_1 is construed in a manner which renders it intuitively unlawlike, the terms "basket b" and "apple" are understood so as to imply finiteness of the class of apples in the basket at time t. Thus, so to speak, the meaning of its constitutive terms alone—without additional factual information—entails that S_1 has a finite scope.—Fundamental laws, then, will have to be construed so as to satisfy what we have called a condition of nonlimited scope; our formulation of that condition however, which refers to what is entailed by "the meaning" of certain expressions, is too vague and will have to be revised later. Let us note in passing that the stipulation here envisaged would bar from the class of fundamental lawlike sentences also such undesirable candidates as "All uranic objects are spherical", where "uranic" means the property

---

in [Counterfactuals], especially part III.—While not concerned with the specific point under discussion, the detailed examination of counterfactual conditionals and their relation to laws of nature, in Chapter VIII of Lewis's work [Analysis], contains important observations on several of the issues raised in the present section.

[22] The view that laws should be construed as not being limited to a finite domain has been expressed, among others, by Popper ([Forschung], section 13) and by Reichenbach ([Logic], p. 369).

[23] [Logic], p. 361.—Our terminology as well as the definitions to be proposed later for the two types of law do not coincide with Reichenbach's, however.

of being the planet Uranus; indeed, while this sentence has universal form, it fails to satisfy the condition of non-limited scope.

In our search for a general characterization of lawlike sentences, we now turn to a second clue which is provided by the sentence S_1. In addition to violating the condition of non-limited scope, this sentence has the peculiarity of making reference to a particular object, the basket b; and this, too, seems to violate the universal character of a law.[24] The restriction which seems indicated here, should however again be applied to fundamental lawlike sentences only; for a true general statement about the free fall of physical bodies on the moon, while referring to a particular object, would still constitute a law, albeit a derivative one.

It seems reasonable to stipulate, therefore, that a fundamental lawlike sentence must be of universal form and must contain no essential—i.e., uneliminable—occurrences of designations for particular objects. But this is not sufficient; indeed, just at this point, a particularly serious difficulty presents itself. Consider the sentence

(S_3) Everything that is either an apple in basket b at time t or a sample of ferric oxide is red.

If we use a special expression, say "x is ferple", as synonymous with "x is either an apple in b at t or a sample of ferric oxide", then the content of S_3 can be expressed in the form

(S_4) Everything that is ferple is red.

The statement thus obtained is of universal form and contains no designations of particular objects, and it also satisfies the condition of non-limited scope; yet clearly, S_4 can qualify as a fundamental lawlike sentence no more than can S_3.

As long as "ferple" is a defined term of our language, the difficulty can readily be met by stipulating that after elimination of defined terms, a fundamental lawlike sentence must not contain essential occurrences of designations for particular objects. But this way out is of no avail when "ferple", or another term of the kind illustrated by it, is a primitive predicate of the language under consideration. This reflection indicates that certain restrictions have to be imposed upon those predicates—i.e., terms for properties or relations,—which may occur in fundamental lawlike sentences.[25]

[24] In physics, the idea that a law should not refer to any particular object has found its expression in the maxim that the general laws of physics should contain no reference to specific space-time points, and that spatio-temporal coördinates should occur in them only in the form of differences or differentials.

[25] The point illustrated by the sentences S_3 and S_4 above was made by Goodman, who has also emphasized the need to impose certain restrictions upon the predicates whose occurrence is to be permissible in lawlike sentences. These predicates are essentially the same as those which Goodman calls projectible. Goodman has suggested that the problems of establishing precise criteria for projectibility, of interpreting counterfactual conditionals, and of defining the concept of law are so intimately related as to be virtually aspects of a single problem. (Cf. his articles [Query] and [Counterfactuals].) One suggestion for an analysis of projectibility has recently been made by Carnap in [Application]. Goodman's note [Infirmities] contains critical observations on Carnap's proposals.

More specifically, the idea suggests itself of permitting a predicate in a fundamental lawlike sentence only if it is purely universal, or, as we shall say, purely qualitative, in character; in other words, if a statement of its meaning does not require reference to any one particular object or spatio-temporal location. Thus, the terms "soft", "green", "warmer than", "as long as", "liquid", "electrically charged", "female", "father of" are purely qualitative predicates, while "taller than the Eiffel Tower", "medieval", "lunar", "arctic", "Ming" are not.[26]

Exclusion from fundamental lawlike sentences of predicates which are not purely qualitative would at the same time ensure satisfaction of the condition of non-limited scope; for the meaning of a purely qualitative predicate does not require a finite extension; and indeed, all the sentences considered above which violate the condition of non-limited scope make explicit or implicit reference to specific objects.

The stipulation just proposed suffers, however, from the vagueness of the concept of purely qualitative predicate. The question whether indication of the meaning of a given predicate in English does or does not require reference to some one specific object does not always permit an unequivocal answer since English as a natural language does not provide explicit definitions or other clear explications of meaning for its terms. It seems therefore reasonable to attempt definition of the concept of law not with respect to English or any other natural language, but rather with respect to a formalized language—let us call it a model language, *L*,—which is governed by a well-determined system of logical rules, and in which every term either is characterized as primitive or is introduced by an explicit definition in terms of the primitives.

This reference to a well-determined system is customary in logical research and is indeed quite natural in the context of any attempt to develop precise criteria for certain logical distinctions. But it does not by itself suffice to overcome the specific difficulty under discussion. For while it is now readily possible to characterize as not purely qualitative all those among the defined predicates in *L* whose definiens contains an essential occurrence of some individual name, our problem remains open for the primitives of the language, whose meanings are not determined by definitions within the language, but rather by semantical rules of interpretation. For we want to permit the interpretation of the primitives of *L* by means of such attributes as blue, hard, solid, warmer, but

[26] That laws, in addition to being of universal form, must contain only purely universal predicates was clearly argued by Popper ([Forschung], sections 14, 15).—Our alternative expression "purely qualitative predicate" was chosen in analogy to Carnap's term "purely qualitative property" (cf. [Application]).—The above characterization of purely universal predicates seems preferable to a simpler and perhaps more customary one, to the effect that a statement of the meaning of the predicate must require no reference to particular objects. For this formulation might be too exclusive since it could be argued that stating the meaning of such purely qualitative terms as "blue" or "hot" requires illustrative reference to some particular object which has the quality in question. The essential point is that no one specific object has to be chosen; any one in the logically unlimited set of blue or of hot objects will do. In explicating the meaning of "taller than the Eiffel Tower", "being an apple in basket b at time t", "medieval", etc., however, reference has to be made to one specific object or to some one in a limited set of objects.

not by the properties of being a descendant of Napoleon, or an arctic animal, or a Greek statue; and the difficulty is precisely that of stating rigorous criteria for the distinction between the permissible and the non-permissible interpretations. Thus the problem of setting up an adequate definition for purely qualitative attributes now arises again; namely for the concepts of the metalanguage in which the semantical interpretation of the primitives is formulated. We may postpone an encounter with the difficulty by presupposing formalization of the semantical meta-language, the meta-meta-language, and so forth; but somewhere, we will have to stop at a non-formalized meta-language, and for it a characterization of purely qualitative predicates will be needed and will present much the same problems as non-formalized English, with which we began. The characterization of a purely qualitative predicate as one whose meaning can be made explicit without reference to any one particular object points to the intended meaning but does not explicate it precisely, and the problem of an adequate definition of purely qualitative predicates remains open.

There can be little doubt, however, that there exists a large number of property and relation terms which would be rather generally recognized as purely qualitative in the sense here pointed out, and as permissible in the formulation of fundamental lawlike sentences; some examples have been given above, and the list could be readily enlarged. When we speak of purely qualitative predicates, we shall henceforth have in mind predicates of this kind.

In the following section, a model language L of a rather simple logical structure will be described, whose primitives will be assumed to be qualitative in the sense just indicated. For this language, the concepts of law and explanation will then be defined in a manner which takes into account the general observations set forth in the present section.

§7. *Definition of law and explanation for a model language.*

Concerning the syntax of our model language L, we make the following assumptions:

L has the syntactical structure of the lower functional calculus without identity sign. In addition to the signs of alternation (disjunction), conjunction, and implication (conditional), and the symbols of universal and existential quantification with respect to individual variables, the vocabulary of L contains individual constants ('a', 'b', $\cdots$), individual variables ('x', 'y', $\cdots$), and predicates of any desired finite degree; the latter may include, in particular, predicates of degree 1 ('P', 'Q', $\cdots$), which express properties of individuals, and predicates of degree 2 ('R', 'S', $\cdots$), which express dyadic relations among individuals.

For simplicity, we assume that all predicates are primitive, i.e., undefined in L, or else that before the criteria subsequently to be developed are applied to a sentence, all defined predicates which it contains are eliminated in favor of primitives.

The syntactical rules for the formation of sentences and for logical inference in L are those of the lower functional calculus. No sentence may contain free variables, so that generality is always expressed by universal quantification.

For later reference, we now define, in purely syntactical terms, a number of auxiliary concepts. In the following definitions, S is always understood to be a sentence in L.

(7.1a) S is formally true (formally false) in L if S (the denial of S) can be proved in L, i.e. by means of the formal rules of logical inference for L. If two sentences are mutually derivable from each other in L, they will be called equivalent.

(7.1b) S is said to be a singular, or alternatively, a molecular sentence if S contains no variables. A singular sentence which contains no statement connectives is also called atomic. Illustrations: The sentences '$R(a, b) \supset (P(a) \cdot \sim Q(a))$', '$\sim Q(a)$', '$(R(a, b)$', '$P(a)$' are all singular, or molecular; the last two are atomic.

(7.1c) S is said to be a generalized sentence if it consists of one or more quantifiers followed by an expression which contains no quantifiers. S is said to be of universal form if it is a generalized sentence and all the quantifiers occurring in it are universal. S is called purely generalized (purely universal) if S is a generalized sentence (is of universal form) and contains no individual constants. S is said to be essentially universal if it is of universal form and not equivalent to a singular sentence. S is called essentially generalized if it is not equivalent to a singular sentence.

Illustrations: '$(x)(P(x) \supset Q(x))$', '$(x)R(a, x)$', '$(x)(P(x) \vee P(a))$', '$(x)(P(x) \vee \sim P(x))$', '$(Ex)(P(x) \cdot \sim Q(x))$', '$(Ex)(y)(R(a, x) \cdot S(a, y))$' are all generalized sentences; the first four are of universal form, the first and fourth are purely universal; the first and second are essentially universal, the third being equivalent to the singular sentence '$P(a)$', and the fourth to '$P(a) \vee \sim P(a)$'. All sentences except the third and fourth are essentially generalized.

Concerning the semantical interpretation of L, we lay down the following two stipulations:

(7.2a) The primitive predicates of L are all purely qualitative.

(7.2b) The universe of discourse of L, i.e., the domain of objects covered by the quantifiers, consists of all physical objects, or of all spatio-temporal locations.

A linguistic framework of the kind here characterized is not sufficient for the formulation of scientific theories since it contains no functors and does not provide the means for dealing with real numbers. Besides, the question is open at present whether a constitution system can be constructed in which all of the concepts of empirical science are reduced, by chains of explicit definitions, to a basis of primitives of a purely qualitative character. Nevertheless, we consider it worthwhile to study the problems at hand for the simplified type of language just described because the analysis of law and explanation is far from trivial even for our model language L, and because that analysis sheds light on the logical character of the concepts under investigation also in their application to more complex contexts.

In accordance with the considerations developed in section 6, we now define:

(7.3a) S is a fundamental lawlike sentence in L if S is purely universal; S is a fundamental law in L if S is purely universal and true.

(7.3b) S is a derivative law in L if (1) S is essentially, but not purely, universal and (2) there exists a set of fundamental laws in L which has S as a consequence.

(7.3c) S is a law in L if it is a fundamental or a derivative law in L.

The fundamental laws as here defined obviously include, besides general statements of empirical character, all those statements of purely universal form which are true on purely logical grounds; i.e. those which are formally true in L, such as '$(x)(P(x) \vee \sim P(x))$', and those whose truth derives exclusively from the interpretation given to its constituents, as is the case with '$(x)(P(x) \supset Q(x))$', if 'P' is interpreted as meaning the property of being a father, and 'Q' that of being male.—The derivative laws, on the other hand, include neither of these categories; indeed, no fundamental law is also a derivative one.

As the primitives of L are purely qualitative, all the statements of universal form in L also satisfy the requirement of non-limited scope, and thus it is readily seen that the concept of law as defined above satisfies all the conditions suggested in section 6.[27]

The explanation of a phenomenon may involve generalized sentences which are not of universal form. We shall use the term "theory" to refer to such sentences, and we define this term by the following chain of definitions:

(7.4a) S is a fundamental theory if S is purely generalized and true.

(7.4b) S is a derivative theory in L if (1) S is essentially, but not purely, generalized and (2) there exists a set of fundamental theories in L which has S as a consequence.

(7.4c) S is a theory in L if it is a fundamental or a derivative theory in L.

By virtue of the above definitions, every law is also a theory, and every theory is true.

With the help of the concepts thus defined, we will now reformulate more precisely our earlier characterization of scientific explanation with specific reference to our model language L. It will be convenient to state our criteria for a sound explanation in the form of a definition for the expression "the ordered couple of sentences, (T, C), constitutes an explanans for the sentence E." Our analysis will be restricted to the explanation of particular events, i.e., to the case where the explanandum, E, is a singular sentence.[28]

[27] As defined above, fundamental laws include universal conditional statements with vacuous antecedents, such as "All mermaids are brunettes". This point does not appear to lead to undesirable consequences in the definition of explanation to be proposed later.—For an illuminating analysis of universal conditionals with vacuous antecedents, see Chapter VIII in Reichenbach's [Logic].

[28] This is not a matter of free choice: The precise rational reconstruction of explanation as applied to general regularities presents peculiar problems for which we can offer no solution at present. The core of the difficulty can be indicated briefly by reference to an example: Kepler's laws, K, may be conjoined with Boyle's law, B, to a stronger law $K.B$; but derivation of K from the latter would not be considered as an explanation of the regularities stated in Kepler's laws; rather, it would be viewed as representing, in effect, a pointless "explanation" of Kepler's laws by themselves. The derivation of Kepler's laws from Newton's laws of motion and of gravitation, on the other hand, would be recognized as a genuine explanation in terms of more comprehensive regularities, or so-called higher-level laws. The problem therefore arises of setting up clear-cut criteria for the distinction of levels of explanation or for a comparison of generalized sentences as to their comprehensiveness. The establishment of adequate criteria for this purpose is as yet an open problem.

In analogy to the concept of lawlike sentence, which need not satisfy a requirement of truth, we will first introduce an auxiliary concept of potential explanans, which is not subject to a requirement of truth; the notion of explanans will then be defined with the help of this auxiliary concept.—The considerations presented in Part I suggest the following initial stipulations:

(7.5) An ordered couple of sentences, (T, C), constitutes a potential explanans for a singular sentence E only if

(1) T is essentially generalized and C is singular

(2) E is derivable in L from T and C jointly, but not from C alone.

(7.6) An ordered couple of sentences, (T, C), constitutes an explanans for a singular sentence E if and only if

(1) (T, C) is a potential explanans for E

(2) T is a theory and C is true.

(7.6) is an explicit definition of explanation in terms of the concept of potential explanation.[29] On the other hand, (7.5) is not suggested as a definition, but as a statement of necessary conditions of potential explanation. These conditions will presently be shown not be sufficient, and additional requirements will be discussed by which (7.5) has to be supplemented in order to provide a definition of potential explanation.

Before we turn to this point, some remarks are called for concerning the formulation of (7.5). The analysis presented in Part I suggests that an explanans for a singular sentence consists of a class of generalized sentences and a class of singular ones. In (7.5), the elements of each of these classes separately are assumed to be conjoined to one sentence. This provision will simplify our formulations, and in the case of generalized sentences, it serves an additional purpose: A class of essentially generalized sentences may be equivalent to a singular sentence; thus, the class {'$P(a)\mathbf{v}(x)Q(x)$', '$P(a)\mathbf{v} \sim (x)Q(x)$'} is equivalent with the sentence '$P(a)$'. Since scientific explanation makes essential use of generalized sentences, sets of laws of this kind have to be ruled out; this is achieved above by combining all the generalized sentences in the explanans into one conjunction, T, and stipulating that T has to be essential generalized. —Again, since scientific explanation makes essential use of generalized sentences, E must not be a consequence of C alone: The law of gravitation, combined with the singular sentence "Mary is blonde and blue-eyed" does not constitute an explanans for "Mary is blonde". The last stipulation in (7.5) introduces the requisite restriction and thus prohibits complete self-explanation of the explanandum, i.e., the derivation of E from some singular sentence which has E as a consequence.—The same restriction also dispenses with the need for a special requirement to the effect that T has to have factual content if (T, C) is to be a potential explanans for an empirical sentence E. For if E is factual, then, since E is a consequence of T and C jointly, but not of C alone, T must be factual, too.

[29] It is necessary to stipulate, in (7.6) (2), that T be a theory rather than merely that T be true, for as was shown in section 6, the generalized sentences occurring in an explanans have to constitute a theory, and not every essentially generalized sentence which is true is actually a theory, i.e., a consequence of a set of purely generalized true sentences.

Our stipulations in (7.5) do not preclude, however, what might be termed partial self-explanation of the explanandum. Consider the sentences T_1 = '$(x)(P(x) \supset Q(x))$', C_1 = '$R(a, b) \cdot P(a) \cdot U(b)$', E_1 = '$Q(a) \cdot R(a, b)$'. They satisfy all the requirements laid down in (7.5), but it seems counterintuitive to say that (T_1, C_1) potentially explains E_1, because the occurrence of the component '$R(a, b)$' of C_1 in the sentence E_1 amounts to a partial explanation of the explanandum by itself. Is it not possible to rule out, by an additional stipulation, all those cases in which E shares part of its content with C, i.e. where C and E have a common consequence which is not formally true in L? This stipulation would be tantamount to the requirement that C and E have to be exhaustive alternatives in the sense that their alternation is formally true, for the content which any two sentences have in common is expressed by their alternation. The proposed restriction, however, would be very severe. For if E does not share even part of its content with C, then C is altogether unnecessary for the derivation of E from T and C, i.e., E can be inferred from T alone. Therefore, in every potential explanation in which the singular component of the explanans is not dispensable, the explanandum is partly explained by itself. Take, for example, the potential explanation of E_2 = '$Q(a)$' by T_2 = '$(x)(P(x) \supset Q(x))$' and C_2 = '$P(a)$', which satisfies (7.5), and which surely is intuitively unobjectionable. Its three components may be equivalently expressed by the following sentences: T_2' = '$(x)(\sim P(x) \vee Q(x))$'; C_2' = '$(P(a) \vee Q(a)) \cdot (P(a) \vee \sim Q(a))$'; E_2' = '$(P(a) \vee Q(a)) \cdot (\sim P(a) \vee Q(a))$'. This reformulation shows that part of the content of the explanandum is contained in the content of the singular component of the explanans and is, in this sense, explained by itself.

Our analysis has reached a point here where the customary intuitive idea of explanation becomes too vague to provide further guidance for rational reconstruction. Indeed, the last illustration strongly suggests that there may be no sharp boundary line which separates the intuitively permissible from the counterintuitive types of partial self-explanation; for even the potential explanation just considered, which is acceptable in its original formulation, might be judged unacceptable on intuitive grounds when transformed into the equivalent version given above.

The point illustrated by the last example is stated more explicitly in the following theorem, which we formulate here without proof.

(7.7) *Theorem*. Let (T, C) be a potential explanans for the singular sentence E. Then there exist three singular sentences, E_1, E_2, and C_1 in L such that E is equivalent to the conjunction $E_1 \cdot E_2$, C is equivalent to the conjunction $C_1 \cdot E_1$, and E_2 can be derived in L from T alone.[30]

In more intuitive terms, this means that if we represent the deductive structure

[30] In the formulation of the above theorem and subsequently, statement connective symbols are used not only as signs *in* L, but also autonymously in speaking *about* compound expressions of L. Thus, when 'S' and 'T' are names or name variables for sentences in L, their conjunction and disjunction will be designated by '$S.T$' and '$S \vee T$', respectively; the conditional which has S as antecedent and T as consequent will be designated by '$S \supset T$', and the denial of S by '$\sim S$'. (Incidentally, this convention has already been used, tacitly, at one place in note 28).

of the given potential explanation by the schema $\{T, C\} \rightarrow E$, then this schema can be restated in the form $\{T, C_1 \cdot E_1\} \rightarrow E_1 \cdot E_2$, where E_2 follows from T alone, so that C_1 is entirely unnecessary as a premise; hence, the deductive schema under consideration can be reduced to $\{T, E_1\} \rightarrow E_1 \cdot E_2$, which can be decomposed into the two deductive schemata $\{T\} \rightarrow E_2$ and $\{E_1\} \rightarrow E_1$. The former of these might be called a purely theoretical explanation of E_2 by T, the latter a complete self-explanation of E_1. Theorem (7.7) shows, in other words, that every explanation whose explanandum is a singular sentence can be decomposed into a purely theoretical explanation and a complete self-explanation; and any explanation of this kind in which the singular constituent of the explanans is not completely unnecessary involves a partial self-explanation of the explanandum.[31]

To prohibit partial self-explanation altogether would therefore mean limitation of explanation to purely theoretical explanation. This measure seems too severely restrictive. On the other hand, an attempt to delimit, by some special rule, the permissible degree of self-explanation does not appear to be warranted because, as we saw, customary usage provides no guidance for such a delimitation, and because no systematic advantage seems to be gained by drawing some arbitrary dividing line. For these reasons, we refrain from laying down stipulations prohibiting partial self-explanation.

The conditions laid down in (7.5) fail to preclude yet another unacceptable type of explanatory argument, which is closely related to complete self-explanation, and which will have to be ruled out by an additional stipulation. The point is, briefly, that if we were to accept (7.5) as a definition, rather than merely as a statement of necessary conditions, for potential explanation, then, as a consequence of (7.6), any given particular fact could be explained by means of any true lawlike sentence whatsoever. More explicitly, if E is a true singular sentence—say, "Mt. Everest is snowcapped",—and T is a law—say, "All metals are good conductors of heat",—then there exists always a true singular sentence C such that E is derivable from T and C, but not from C alone; in other words, such that (7.5) is satisfied. Indeed, let T_s be some arbitrarily chosen particular instance of T, such as "If the Eiffel Tower is metal, it is a good conductor of heat". Now since E is true, so is the conditional $T_s \supset E$, and if the latter is chosen as the sentence C, then T, C, E satisfy the conditions laid down in (7.5).

In order to isolate the distinctive characteristic of this specious type of explanation, let us examine an especially simple case of the objectionable kind.

[31] The characteristic here referred to as partial self-explanation has to be distinguished from what is sometimes called the circularity of scientific explanation. The latter phrase has been used to cover two entirely different ideas. (a) One of these is the contention that the explanatory principles adduced in accounting for a specific phenomenon are inferred from that phenomenon, so that the entire explanatory process is circular. This belief is false, since general laws cannot be inferred from singular sentences. (b) It has also been argued that in a sound explanation the content of the explanandum is contained in that of the explanans. That is correct since the explanandum is a logical consequence of the explanans; but this peculiarity does not make scientific explanation trivially circular since the general laws occurring in the explanans go far beyond the content of the specific explanandum. For a fuller discussion of the circularity objection, see Feigl, [Operationism], pp. 286 ff, where this issue is dealt with very clearly.

Let T_1 = '$(x)P(x)$' and E_1 = '$R(a, b)$'; then the sentence C_1 = '$P(a) \supset R(a, b)$' is formed in accordance with the preceding instructions, and T_1, C_1, E_1 satisfy the conditions (7.5). Yet, as the preceding example illustrates, we would not say that (T_1, C_1) constitutes a potential explanans for E_1. The rationale for the verdict may be stated as follows: If the theory T_1 on which the explanation rests, is actually true, then the sentence C_1, which can also be put into the form '$\sim P(a) \vee R(a, b)$', can be verified, or shown to be true, only by verifying '$R(a, b)$', i.e., E_1. In this broader sense, E_1 is here explained by itself. And indeed, the peculiarity just pointed out clearly deprives the proposed potential explanation for E_1 of the predictive import which, as was noted in Part I, is essential for scientific explanation: E_1 could not possibly be predicted on the basis of T_1 and C_1 since the truth of C_1 cannot be ascertained in any manner which does not include verification of E_1. (7.5) should therefore be supplemented by a stipulation to the effect that if (T, C) is to be a potential explanans for E, then the assumption that T is true must not imply that verification of C necessitates verification of E.[32]

How can this idea be stated more precisely? Study of an illustration will suggest a definition of verification for molecular sentences. The sentence M = '$(\sim P(a) \cdot Q(a)) \vee R(a, b)$' may be verified in two different ways, either by ascertaining the truth of the two sentences '$\sim P(a)$' and '$Q(a)$', which jointly have M as a consequence, or by establishing the truth of the sentence '$R(a, b)$', which, again, has M as a consequence. Let us say that S is a basic sentence in L if S is either an atomic sentence or the denial of an atomic sentence in L. Verification of a molecular sentence S may then be defined generally as establishment of the truth of some class of basic sentences which has S as a consequence. Hence, the intended additional stipulation may be restated: The assumption that T is true must not imply that every class of true basic sentences which has C as a consequence also has E as a consequence.

As brief reflection shows, this stipulation may be expressed in the following form, which avoids reference to truth: T must be compatible in L with at least one class of basic sentences which has C but not E as a consequence; or, equivalently: There must exist at least one class of basic sentences which has C, but neither $\sim T$ nor E as a consequence in L.

If this requirement is met, then surely E cannot be a consequence of C, for otherwise there could be no class of basic sentences which has C but not E as a consequence; hence, supplementation of (7.5) by the new condition renders the second stipulation in (7.5) (2) superfluous.—We now define potential explanation as follows:

(7.8) An ordered couple of sentences, (T, C), constitutes a potential explanans for a singular sentence E if and only if the following conditions are satisfied:

(1) T is essentially generalized and C is singular

[32] It is important to distinguish clearly between the following two cases: (a) If T is true then C cannot be true without E being true; and (b) If T is true, C cannot be verified without E being verified.—Condition (a) must be satisfied by any potential explanation; the much more restictive condition (b) must not be satisfied if (T,C) is to be a potential explanans for E.

(2) E is derivable in L from T and C jointly
(3) T is compatible with at least one class of basic sentences which has C but not E as a consequence.

The definition of the concept of explanans by means of that of potential explanans as formulated in (7.6) remains unchanged.

In terms of our concept of explanans, we can give the following interpretation to the frequently used phrase "this fact is explainable by means of that theory":

(7.9) A singular sentence E is explainable by a theory T if there exists a singular sentence C such that (T, C) constitutes an explanans for E.

The concept of causal explanation, which has been examined here, is capable of various generalizations. One of these consists in permitting T to include statistical laws. This requires, however, a previous strengthening of the means of expression available in L, or the use of a complex theoretical apparatus in the metalanguage.—On the other hand, and independently of the admission of statistical laws among the explanatory principles, we may replace the strictly deductive requirement that E has to be a consequence of T and C jointly by the more liberal inductive one that E has to have a high degree of confirmation relatively to the conjunction of T and C. Both of these extensions of the concept of explanation open important prospects and raise a variety of new problems. In the present essay, however, these issues will not be further pursued.

PART IV. THE SYSTEMATIC POWER OF A THEORY

§8. *Explication of the concept of systematic power.*

Scientific laws and theories have the function of establishing systematic connections among the data of our experience, so as to make possible the derivation of some of those data from others. According as, at the time of the derivation, the derived data are, or are not yet, known to have occurred, the derivation is referred to as explanation or as prediction. Now it seems sometimes possible to compare different theories, at least in an intuitive manner, in regard to their explanatory or predictive powers: Some theories seem powerful in the sense of permitting the derivation of many data from a small amount of initial information, others seem less powerful, demanding comparatively more initial data, or yielding fewer results. Is it possible to give a precise interpretation to comparisons of this kind by defining, in a completely general manner, a numerical measure for the explanatory or predictive power of a theory? In the present section, we shall develop such a definition and examine some of its implications; in the following section, the definition will be expanded and a general theory of the concept under consideration will be outlined.

Since explanation and prediction have the same logical structure, namely that of a deductive systematization, we shall use the neutral term "systematic power" to refer to the intended concept. As is suggested by the preceding intuitive characterization, the systematic power of a theory T will be reflected in the ratio of the amount of information derivable by means of T to the amount of initial information required for that derivation. This ratio will obviously depend on the particular set of data, or of information, to which T is applied, and we shall therefore relativize our concept accordingly. Our aim, then, is to construct a definition for $s(T, K)$, the systematic power of a theory T with respect to a finite class K of data, or the degree to which T deductively systematizes the information contained in K.

Our concepts will be constructed again with specific reference to the language L. Any singular sentence in L will be said to express a potential datum, and K will accordingly be construed as a finite class of singular sentences[33]. T will be construed in a much broader sense than in the preceding sections; it may be any sentence in L, no matter whether essentially generalized or not. This liberal convention is adopted in the interest of the generality and simplicity of the definitions and theorems now to be developed.

To obtain values between 0 and 1 inclusive, we might now try to identify $s(T, K)$ with the percentage of those sentences in K which are derivable from the remainder by means of T. Thus, if $K_1 = \{`P(a)', `Q(a)', `\sim P(b)', `\sim Q(b)', `Q(c)', `\sim P(d)'\}$, and $T_1 = `(x)(P(x) \supset Q(x))'$, then exactly the second and third sentence in $K_1$ are derivable by means of $T_1$ from the remainder, in fact from the first and fourth sentence. We might therefore consider setting $s(T_1, K_1) = 2/6 = 1/3$. But then, for the class $K_2 = \{`P(a) \cdot Q(a)', `\sim P(b) \cdot \sim Q(b)', `Q(c)', `\sim P(d)'\}$, the same $T_1$ would have the $s$-value 0, although $K_2$ contains exactly the same information as $K_1$; again, for yet another formulation of that information, namely, $K_3 = \{`P(a) \cdot \sim Q(b)', `Q(a) \cdot \sim P(b)', `Q(c)', `\sim P(d)'\}$, T_1 would have the s-value 1/4, and so on. But what we seek is a measure of the degree to which a given theory deductively systematizes a given body of factual information, i.e., a certain content, irrespective of the particular structure and grouping of the sentences in which that content happens to be expressed. We shall therefore make use of a method which represents the contents of any singular sentence or class of singular sentences as composed of certain uniquely determined smallest bits of information. By applying our general idea to these bits, we shall obtain a measure for the systematic power of T in K which is independent of the way in which the content of K is formulated. The sentences expressing those smallest bits of information will be called minimal sentences, and an exact formulation of the proposed procedure will be made possible by an explicit definition of this auxiliary concept. To this point we now turn.

If, as will be assumed here, the vocabulary of L contains fixed finite numbers of individual constants and of predicate constants, then only a certain finite number, say n, of different atomic sentences can be formulated in L. By a minimal

[33] As this stipulation shows, the term "datum" is here understood as covering actual as well as potential data. The convention that any singular sentence expresses a potential datum is plausible especially if the primitive predicates of L refer to attributes whose presence absence in specific instances can be ascertained by direct observation. In this case, each singular sentence in L may be considered as expressing a potential datum, in the sense of describing a logically possible state of affairs whose existence might be ascertained by direct observation.—The assumption that the primitives of L express directly observable attributes is, however, not essential for the definition and the formal theory of systematic power set forth in sections 8 and 9.

sentence in L, we will understand a disjunction of any number k ($0 \leq k \leq n$) of different atomic sentences and the denials of the n-k remaining ones. Clearly, n atomic sentences determine 2^n minimal sentences. Thus, if a language L_1 contains exactly one individual constant, 'a', and exactly two primitive predicates, 'P' and 'Q', both of degree 1, then L_1 contains two atomic sentences, '$P(a)$' and '$Q(a)$', and four minimal sentences, namely, '$P(a) \vee Q(a)$', '$P(a) \vee \sim Q(a)$', '$\sim P(a) \vee Q(a)$', '$\sim P(a) \vee \sim Q(a)$'. If another language, L_2, contains in addition to the vocabulary of L_1 a second individual constant, 'b', and a predicate 'R' of degree 2, then L_2 contains eight atomic sentences and 256 minimal sentences, such as '$P(a) \vee P(b) \vee \sim Q(a) \vee Q(b) \vee R(a, a) \vee R(a, b) \vee \sim R(b, a) \vee \sim R(b, b)$'.

The term "minimal sentence" is to indicate that the statements in question are the singular sentences of smallest non-zero content in L, which means that every singular sentence in L which follows from a minimal sentence is either equivalent with that minimal sentence or formally true in L. However, minimal sentences do have consequences other than themselves which are not formally true in L, but these are not of singular form; '$(Ex)(P(x) \vee Q(x))$' is such a consequence of '$P(a) \vee Q(a)$' in L_1 above.

Furthermore, no two minimal sentences have any consequence in common which is not formally true in L; in other words, the contents of any two minimal sentences are mutually exclusive.

By virtue of the principles of the sentential calculus, every singular sentence which is not formally true in L can be transformed into a conjunction of uniquely determined minimal sentences; this conjunction will be called the minimal normal form of the sentence. Thus, e.g., in the language L_1 referred to above, the sentences '$P(a)$' and '$Q(a)$' have the minimal normal forms '$P(a) \vee Q(a)) \cdot (P(a) \vee \sim Q(a))$', and '$(P(a) \vee Q(a)) \cdot (\sim P(a) \vee Q(a))$', respectively; in L_2, the same sentences have minimal normal forms consisting of 128 conjoined minimal sentences each.—If a sentence is formally true in L, its content is zero, and it cannot be represented by a conjunction of minimal sentences. It will be convenient, however, to say that the minimal normal form of a formally true sentence in L is the vacuous conjunction of minimal sentences, which does not contain a single term.

As a consequence of the principle just mentioned, any class of singular sentences which are not all formally true can be represented by a sentence in minimal normal form. The basic idea outlined above for the explication of the concept of systematic power can now be expressed by the following definition:

(8.1) Let T be any sentence in L, and K any finite class of singular sentences in L which are not all formally true. If K' is the class of minimal sentences which occur in the minimal normal form of K, consider all divisions of K' into two mutually exclusive subclasses, K'_1 and K'_2, such that every sentence in K'_2 is derivable from K'_1 by means of T. Each division of this kind determines a ratio $n(K'_2)/n(K')$, i.e. the number of minimal sentences in K'_2 divided by the total number of minimal sentences in K'. Among the values of these ratios,

there must be a largest one; $s(T, K)$ is to equal that maximum ratio. (Note that if all the elements of K were formally true, $n(K')$ would be 0 and the above ratio would not be defined.)

Illustration: Let L_1 contain only one individual constant, 'a', and only two predicates, 'P' and 'Q', both of degree 1. In L_1, let T = '$(x)(P(x) \supset Q(x))$', $K = \{$'$P(a)$', '$Q(a)$'$\}$. Then we have $K' = \{$'$P(a) \vee Q(a)$', '$P(a) \vee \sim Q(a)$', '$\sim P(a) \vee Q(a)$'$\}$. From the subclass K_1' consisting of the first two elements of K'—which together are equivalent to '$P(a)$'—we can derive, by means of T, the sentence '$Q(a)$', and from it, by pure logic, the third element of K'; it constitutes the only element of K_2'. No "better" systematization is possible, hence $s(T, K) = 1/3$.

Our definition leaves open, and is independent of, the question whether for a given K' there might not exist different divisions each of which would yield the maximum value for $n(K_2')/n(K')$. Actually, this can never happen: there exists always exactly one optimal subdivision of a given K'. This fact is a corollary of a general theorem, to which we now turn. It will be noticed that in the last illustration, K_2' can be derived from T alone, without the use of K_1' as a premise; indeed, '$\sim P(a) \vee Q(a)$' is but a substitution instance of the sentence '$(x)(\sim P(x) \vee Q(x))$', which is equivalent to T. The theorem now to be formulated, which might appear surprising at first, shows that this observation applies analogously in all other cases.

(8.2) Theorem. Let T be any sentence, K' a class of minimal sentences, and K_2' a subclass of K' such that every sentence in K_2' is derivable by means of T from the class $K - K_2'$; then every sentence in K_2' is derivable from T alone.

The proof, in outline, is as follows: Since the contents of any two different minimal sentences are mutually exclusive, so must be the contents of K_1' and K_2', which have not a single minimal sentence in common. But since the sentences of K_2' follow from K_1' and T jointly, they must therefore follow from T alone.

We note the following consequences of our theorem:

(8.2a) Theorem. In any class K' of minimal sentences, the largest subclass which is derivable from the remainder by means of a sentence T is identical with the class of those elements in K' which are derivable from T alone.

(8.2b) Theorem. Let T be any sentence, K a class of singular sentences which are not all formally true, K' the equivalent class of minimal sentences, and K_t' the class of those among the latter which are derivable from T alone. Then the concept s defined in (8.1) satisfies the following equation:

$$s(T, K) = n(K_t')/n(K')$$

§9. *Systematic power and logical probability of a theory. Generalization of the concept of systematic power.*

The concept of systematic power is closely related to that of degree of confirmation, or logical probability, of a theory. A study of this relationship will shed new light on the proposed definition of s, will suggest certain ways of generalizing it, and will finally lead to a general theory of systematic power which is formally analogous to that of logical probability.

The concept of logical probability, or degree of confirmation, is the central concept of inductive logic. Recently, different explicit definitions for this concept have been proposed, for languages of a structure similar to that of our model language, by Carnap[34] and by Helmer, Hempel, and Oppenheim[35].

While the definition of s proposed in the preceding section rests on the concept of minimal sentence, the basic concept in the construction of a measure for logical probability is that of state description, or, as we shall also say, of maximal sentence. A maximal sentence is the dual[36] of a minimal sentence in L; it is a conjunction of k ($0 \leqq k \leqq n$) different atomic sentences and of the denials of the remaining n-k atomic sentences. In a language with n atomic sentences, there exist 2^n state descriptions. Thus, e.g., the language L_1 repeatedly mentioned in §8 contains the following four maximal sentences: '$P(a) \cdot Q(a)$', '$P(a) \cdot \sim Q(a)$', '$\sim P(a) \cdot Q(a)$', '$\sim P(a) \cdot \sim Q(a)$'.

The term "maximal sentence" is to indicate that the sentences in question are the singular sentences of maximum non-universal content in L, which means that every singular sentence in L which has a maximal sentence as a consequence is either equivalent with that maximal sentence or formally false in L.

As we saw, every singular sentence can be represented in a conjunctive, or minimal, normal form, i.e., as a conjunction of certain uniquely determined minimal sentences; similarly, every singular sentence can be expressed also in a disjunctive, or maximal, normal form, i.e. as a disjunction of certain uniquely determined maximal sentences. In the language L_1, for example, '$P(a)$' has the minimal normal form '$(P(a) \mathbf{v} Q(a)) \cdot (P(a) \mathbf{v} \sim Q(a))$' and the maximal normal form '$(P(a) \cdot Q(a)) \mathbf{v} (P(a) \cdot \sim Q(a))$'; the sentence '$P(a) \supset Q(a)$' has the minimal normal form '$\sim P(a) \mathbf{v} Q(a)$' and the maximal normal form '$(P(a) \cdot Q(a)) \mathbf{v} (\sim P(a) \cdot Q(a)) \mathbf{v} (\sim P(a) \cdot \sim Q(a))$'; the minimal normal form of a formally true sentence is the vacuous conjunction, while its maximal normal form is the disjunction of all four state descriptions in L_1. The minimal normal form of any formally false sentence is the conjunction of all four minimal sentences in L_1, while its maximal normal form is the vacuous disjunction, as we shall say.

The minimal normal form of a singular sentence is well suited as an indicator of its content, for it represents the sentence as a conjunction of standard components whose contents are minimal and mutually exclusive. The maximal normal form of a sentence is suited as an indicator of its range, that is, intuitively speaking, of the variety of its different possible realizations, or of the variety of

[34] Cf. especially [Inductive Logic], [Concepts], [Application].

[35] See Helmer and Oppenheim, [Probability]; Hempel and Oppenheim, [Degree].—Certain general aspects of the relationship between the confirmation of a theory and its predictive or systematic success are examined in Hempel, [Studies], Part II, sections 7 and 8. The definition of s developed in the present essay establishes a quantitative counterpart of what, in that paper, is characterized, in non-numerical terms, as the prediction criterion of confirmation.

[36] For a definition and discussion of this concept, cf. Church, [Logic], p. 172.

those possible states of the world which, if realized, would make the statement true. Indeed, each maximal sentence may be considered as describing, as completely as is possible in L, one possible state of the world; and the state descriptions constituting the maximal normal form of a given singular sentence simply list those among the possible states which would make the sentence true.

Just as the contents of any two different minimal sentences, so also the ranges of any two maximal sentences are mutually exclusive: No possible state of the world can make two different maximal sentences true because any two maximal sentences are obviously incompatible with each other.[37]

Range and content of a sentence vary inversely. The more a sentence asserts, the smaller the variety of its possible realizations, and conversely. This relationship is reflected in the fact that the larger the number of constituents in the minimal normal form of a singular sentence, the smaller the number of constituents in its maximal normal form, and conversely. In fact, if the minimal normal form of a singular sentence U contains m_U of the $m = 2^n$ minimal sentences in L, then its maximal normal form contains $l_U = m - m_U$ of the m maximal sentences in L. This is illustrated by our last four examples, where $m = 4$, and $m_U = 2, 1, 0, 4$ respectively.

The preceding observations suggest that the content of any singular sentence U might be measured by the corresponding number m_U or by some magnitude proportional to it. Now it will prove convenient to restrict the values of the content measure function to the interval from 0 to 1, inclusive; and therefore, we define a measure, $g_1(U)$, for the content of any singular sentence in L by the formula

$$g_1(U) = m_U/m \tag{9.1}$$

To any finite class K of singular sentences, we assign, as a measure $g_1(K)$ of its content, the value $g_1(S)$, where S is the conjunction of the elements of K.

By virtue of this definition, the equation in theorem (8.2b) may be rewritten:

$$s(T, K) = g_1(K'_t)/g_1(K')$$

Here, K'_t is the class of all those minimal sentences in K' which are consequences of T. In the special case where T is a singular sentence, K'_t is therefore equivalent with $T \mathbf{v} S$, where S is the conjunction of all the elements of K'. Hence, the preceding equation may then be transformed into

$$s(T, S) = g_1(T \mathbf{v} S)/g_1(S) \tag{9.2}$$

This formula holds when T and S are singular sentences, and S is not formally true. It bears a striking resemblance to the general schema for the definition of the logical probability of T in regard to S:

$$p(T, S) = r(T \cdot S)/r(S) \tag{9.3}$$

[37] A more detailed discussion of the concept of range may be found in Carnap, [Inductive Logic], section 2, and in Carnap, [Semantics], sections 18 and 19, where the relation of range and content is examined at length.

Here, $r(U)$ is, for any sentence U in L, a measure of the range of U, T is any sentence in L, and S any sentence in L with $r(S) \neq 0$.

The several specific definitions which have been proposed for the concept of logical probability accord essentially with the pattern exhibited by (9.3)[38], but they differ in their choice of a specific measure function for ranges, i.e. in their definition of r. One idea which comes to mind is to assign, to any singular sentence U whose maximal normal form contains l_U maximal sentences, the range measure

$$r_1(U) = l_U/m \tag{9.4}$$

which obviously is defined in strict analogy to the content measure g_1 for singular sentences as introduced in (9.1). For every singular sentence U, the two measures add up to unity:

$$r_1(U) + g_1(U) = (l_U + m_U)/m = 1 \tag{9.5}$$

As Carnap has shown, however, the range measure r_1 confers upon the corresponding concept of logical probability, i.e., upon the concept p_1 defined by means of it according to the schema (9.3), certain characteristics which are incompatible with the intended meaning of logical probability[39]; and Carnap as well as Helmer jointly with the present authors have suggested certain alternative measure functions for ranges, which lead to more satisfactory concepts of probability or of degree of confirmation. While we need not enter into details here, the following general remarks seem indicated to prepare the subsequent discussion.

The function r_1 measures the range of a singular sentence essentially by counting the number of maximal sentences in its maximal normal form; it thus gives equal weight to all maximal sentences (definition (9.1) deals analogously with minimal sentences). The alternative definitions just referred to are based on a different procedure. Carnap, in particular, lays down a rule which assigns a specific weight, i.e. a specific value of r, to each maximal sentence, but these weights are not the same for all maximal sentences. He then defines the range measure of any other singular sentence as the sum of the measures of its constituent maximal sentences. In terms of the function thus obtained—let us call it r_2—Carnap defines the corresponding concept of logical probability, which we shall call p_2, for singular sentences T, S in accordance with the schema (9.3): $p_2(T, S) = r_2(T. S)/r_2(S)$. The definitions of r_2 and p_2 are then extended, by means of certain limiting processes, to the cases where T and S are no longer both singular.[40]

[38] In Carnap's theory of logical probability, $p(T, S)$ is defined, for certain cases, as the limit which the function $r(T. S)/r(S)$ assumes under specified conditions (cf. Carnap, [Inductive Logic], p. 75); but we shall refrain here from considering this generalization of that type of definition which is represented by (9.3).

[39] [Inductive Logic], pp. 80–81.

[40] The alternative approach suggested by Helmer and the present authors involves use of a range measure function r_I which depends in a specified manner on the empirical information I available; hence, the range measure of any sentence U is determined only if a sentence

Now it can readily be seen that just as the function r_1 defined in (9.5) is but but one among an infinity of possible range measures, so the analogous function g_1 defined in (9.1) is but one among an infinity of possible content measures; and just as each range measure may serve to define, according to the schema (9.3), a corresponding measure of logical probability, so each content measure function may serve to define, by means of the schema illustrated by (9.2), a corresponding measure of systematic power. The method which suggests itself here for obtaining alternative content measure functions is to choose some range measure r other than r_1 and then to *define* a corresponding content measure g in terms of it by means of the formula

$$(9.6) \qquad g(U) = 1 - r(U)$$

so that g and r satisfy the analogue to (9.5) by definition. The function g thus defined will lead in turn, via a definition analogous to (9.2), to a corresponding concept s. Let us now consider this procedure a little more closely.

We assume that a function r is given which satisfies the customary requirements for range measures, namely:

(9.7) 1. $r(U)$ is uniquely determined for *all* sentences U in L.

2. $0 \leqq r(U) \leqq 1$ for every sentence U in L.
3. $r(U) = 1$ if the sentence U is formally true in L and thus has universal range.
4. $r(U_1 \vee U_2) = r(U_1) + r(U_2)$ for any two sentences U_1, U_2 whose ranges are mutually exclusive, i.e., whose conjunction is formally false.

In terms of the given range measure let the corresponding content measure g be defined by means of (9.6). Then g can readily be shown to satisfy the following conditions:

(9.8) 1. $g(U)$ is uniquely determined for *all* sentences U in L.

2. $0 \leqq g(U) \leqq 1$ for every sentence U in L.
3. $g(U) = 1$ if the sentence U is formally false in L and thus has universal content.

I, expressing the available empirical information, is given. In terms of this range measure function, the concept of degree of confirmation, dc, can be defined by means of a formula similar to (9.3). The value of $dc(T, S)$ is not defined, however, in certain cases where S is generalized, as has been pointed out by McKinsey (cf. [Review]); also, the concept dc does not satisfy all the theorems of elementary probability theory (cf. the discussion of this point in the first two articles mentioned in note (35)); therefore, the degree of confirmation of a theory relatively to a given evidence is not a probability in the strict sense of the word. On the other hand, the definition of dc here referred to has certain methodologically desirable features, and it might therefore be of interest to construct a related concept of systematic power by means of the range measure function r_I. In the present paper, however, this question will not be pursued.

4. $g(U_1 \cdot U_2) = g(U_1) + g(U_2)$ for any two sentences U_1, U_2 whose contents are mutually exclusive, i.e., whose disjunction is formally true.

In analogy to (9.2), we next define, by means of g, a corresponding function s:

$$s(T, S) = g(T \vee S)/g(S) \quad (9.9)$$

This function is determined for every sentence T, and for every sentence S with $g(S) \neq 0$, whereas the definition of systematic power given in §8 was restricted to those cases where S is singular and not formally true. Finally, our range measure r determines a corresponding probability function by virtue of the definition

$$p(T, S) = r(T \cdot S)/r(S) \quad (9.10)$$

This formula determines the function p for any sentence T, and for any sentence S with $r(S) \neq 0$.

In this manner, every range measure r which satisfies (9.7) determines uniquely a corresponding content measure g which satisfies (9.8), a corresponding function s, defined by (9.9), and a corresponding function p, defined by (9.10). As a consequence of (9.7) and (9.10), the function p can be shown to satisfy the elementary laws of probability theory, especially those listed in (9.12) below; and by virtue of these, it is finally possible to establish a very simple relationship which obtains, for any given range measure r, between the corresponding concepts $p(T, S)$ and $s(T, S)$. Indeed, we have

$$(9.11) \quad \begin{aligned} s(T, S) &= g(T \vee S)/g(S) \\ &= (1 - r(T \vee S))/(1 - r(S)) \\ &= r(\sim (T \vee S))/r(\sim S) \\ &= r(\sim T \cdot \sim S)/r(\sim S) \\ &= p(\sim T, \sim S) \end{aligned}$$

We now list, without proof, some theorems concerning p and s which follow from our assumptions and definitions; they hold in all cases where the values of p and s referred to exist, i.e., where the r-value of the second arguments of p, and the g-value of the second arguments of s, is not 0.

(9.12) (1) a. $0 \leqq p(T, S) \leqq 1$

b. $0 \leqq s(T, S) \leqq 1$

(2) a. $p(\sim T, S) = 1 - p(T, S)$

b. $s(\sim T, S) = 1 - s(T, S)$

(3) a. $p(T_1 \vee T_2, S) = p(T_1, S) + p(T_2, S) - p(T_1 \cdot T_2, S)$

b. $s(T_1 \cdot T_2, S) = s(T_1, S) + s(T_2, S) - s(T_1 \vee T_2, S)$

(4) a. $p(T_1 \cdot T_2, S) = p(T_1, S) \cdot p(T_2, T_1 \cdot S)$

b. $s(T_1 \vee T_2, S) = s(T_1, S) \cdot s(T_2, T_1 \vee S)$

In the above grouping, these theorems exemplify the relationship of dual correspondence which obtains between p and s. A general characterization of this correspondence is given in the following theorem, which can be proved on the basis of (9.11), and which is stated here in a slightly informal manner in order to avoid the tedium of lengthy formulations.

(9.13) *Dualism theorem.* From any demonstrable general formula expressing an equality or an inequality concerning p, a demonstrable formula concerning s is obtained if 'p' is replaced, throughout, by 's', and '$\cdot$' and '$\mathbf{v}$' are exchanged for each other. The same exchange, and replacement of 's' by 'p', conversely transforms any theorem expressing an equality or an inequality concerning s into a theorem about p.

We began our analysis of the systematic power of a theory in regard to a class of data by interpreting this concept, in §8, as a measure of the optimum ratio of those among the given data which are derivable from the remainder by means of the theory. Systematic elaboration of this idea has led to the definition, in the present section, of a more general concept of systematic power, which proved to be the dual counterpart of the concept of logical probability. This extension of our original interpretation yields a simpler and more comprehensive theory than would have been attainable on the basis of our initial definition.

But the theory of systematic power, in its narrower as well as in its generalized version, is, just like the theory of logical probability, purely formal in character, and a significant application of either theory in epistemology or the methodology of science requires the solution of certain fundamental problems which concern the logical structure of the language of science and the interpretation of its concepts. One urgent desideratum here is the further elucidation of the requirement of purely qualitative primitives in the language of science; another crucial problem is that of choosing, among an infinity of formal possibilities, an adequate range measure r. The complexity and difficulty of the issues which arise in these contexts has been brought to light by recent investigations[41]; it can only be hoped that recent developments in formal theory will soon be followed by progress in solving those open problems and thus clarifying the conditions for a sound application of the theories of logical probability and of systematic power

Queens College, Flushing, N. Y.
Princeton, N. J.

BIBLIOGRAPHY

Throughout the article, the abbreviated titles in brackets are used for reference

Beard, Charles A., and Hook, Sidney. [Terminology] Problems of terminology in historical writing. Chapter IV of Theory and practice in historical study: A report of the Committee on Historiography. Social Science Research Council, New York, 1946.

Bergmann, Gustav. [Emergence] Holism, historicism, and emergence. *Philosophy of Science*, vol. 11 (1944), pp. 209–221.

Bonfante, G. [Semantics] Semantics, language. An article in P. L. Harriman, ed., The encyclopedia of psychology. Philosophical Library, New York, 1946.

Broad, C. D. [Mind] The mind and its place in nature. New York, 1925.

[41] Cf. especially Goodman, [Query], [Counterfactuals], [Infirmities], and Carnap, [Application]. See also notes (21) and (25).

Carnap, Rudolf. [Semantics] Introduction to semantics. Harvard University Press, 1942.

———. [Inductive Logic] On inductive logic. *Philosophy of science*, vol 12 (1945), pp. 72–97.

———. [Concepts] The two concepts of probability. *Philosophy and phenomenological research*, vol. 5 (1945), pp. 513–532.

———. [Remarks] Remarks on induction and truth. *Philosophy and phenomenological research*, vol. 6 (1946), pp. 590–602.

———. [Application] On the application of inductive logic. *Philosophy and phenomenological research*, vol. 8 (1947), pp. 133–147.

Chisholm, Roderick M. [Conditional] The contrary-to-fact conditional. *Mind*, vol. 55 (1946), pp. 289–307.

Church, Alonzo. [Logic] Logic, formal. An article in Dagobert D. Runes, ed. The dictionary of philosophy. Philosophical Library, New York, 1942.

Ducasse, C. J. [Explanation] Explanation, mechanism, and teleology. *The journal of philosophy*, vol. 22 (1925), pp. 150–155.

Feigl, Herbert. [Operationism] Operationism and scientific method. *Psychological review*, vol. 52 (1945), pp. 250–259 and 284–288.

Goodman, Nelson. [Query] A query on confirmation. *The journal of philosophy*, vol. 43 (1946), pp. 383–385.

———. [Counterfactuals]. The problem of counterfactual conditionals. *The journal of philosophy*, vol. 44 (1947), pp. 113–128.

———. [Infirmities] On infirmities of confirmation theory. *Philosophy and phenomenological research*, vol. 8 (1947), pp. 149–151.

Grelling, Kurt and Oppenheim, Paul. [Gestaltbegriff] Der Gestaltbegriff im Lichte der neuen Logik. *Erkenntnis*, vol. 7 (1937–38), pp. 211–225 and 357–359.

Grelling, Kurt and Oppenheim, Paul. [Functional Whole] Logical Analysis of "Gestalt" as "Functional whole". Preprinted for distribution at Fifth Internat. Congress for the Unity of Science, Cambridge, Mass., 1939.

Helmer, Olaf and Oppenheim, Paul. [Probability] A syntactical definition of probability and of degree of confirmation. *The journal of symbolic logic*, vol. 10 (1945), pp. 25–60.

Hempel, Carl G. [Laws] The function of general laws in history. *The journal of philosophy*, vol. 39 (1942), pp. 35–48.

———. [Studies] Studies in the logic of confirmation. *Mind*, vol. 54 (1945); Part I: pp. 1–26, Part II: pp. 97–121.

Hempel, Carl G. and Oppenheim, Paul. [Degree] A definition of "degree of confirmation". *Philosophy of science*, vol. 12 (1945), pp. 98–115.

Henle, Paul. [Emergence] The status of emergence. *The journal of philosophy*, vol. 39 (1942), pp. 486–493.

Hospers, John. [Explanation] On explanation. *The journal of philosophy*, vol. 43 (1946), pp. 337–356.

Hull, Clark L. [Variables] The problem of intervening variables in molar behavior theory. *Psychological review*, vol. 50 (1943), pp. 273–291.

———. [Principles] Principles of behavior. New York, 1943.

Jevons, W. Stanley. [Principles] The principles of science. London, 1924. (1st ed. 1874).

Kaufmann, Felix. [Methodology] Methodology of the social sciences. New York, 1944.

Knight, Frank H. [Limitations] The limitations of scientific method in economics. In Tugwell, R., ed., The trend of economics. New York, 1924.

Koch, Sigmund. [Motivation] The logical character of the motivation concept. *Psychological review*, vol. 48 (1941). Part I: pp. 15–38, Part II: pp. 127–154.

Langford, C. H. [Review] Review in *The journal of symbolic logic*, vol. 6 (1941), pp. 67–68

Lewis, C. I. [Analysis] An analysis of knowledge and valuation. La Salle, Ill., 1946.

McKinsey, J. C. C. [Review] Review of Helmer and Oppenheim, [Probability]. *Mathematical reviews*, vol. 7 (1946), p. 45.

Mill, John Stuart. [Logic] A system of Logic.
Morgan, C. Lloyd. Emergent evolution, New York, 1923.
———. The emergence of novelty. New York, 1933.
Popper, Karl. [Forschung] Logik der Forschung. Wien, 1935.
———. [Society] The open society and its enemies. London, 1945.
Reichenbach, Hans. [Logic] Elements of symbolic logic. New York, 1947.
———. [Quantum mechanics] Philosophic foundations of quantum mechanics. University of California Press, 1944.
Rosenblueth, A., Wiener, N., and Bigelow, J. [Teleology] Behavior, Purpose, and Teleology. *Philosophy of science*, vol. 10 (1943), pp. 18–24.
Stace, W. T. [Novelty] Novelty, indeterminism and emergence. *Philosophical review*, vol. 48 (1939), pp. 296–310.
Tarski, Alfred. [Truth] The semantical conception of truth, and the foundations of semantics. *Philosophy and phenomenological research*, vol. 4 (1944), pp. 341–376.
Tolman, Edward Chase. [Behavior] Purposive behavior in animals and men. New York 1932.
White, Morton G. [Explanation] Historical explanation. *Mind*, vol. 52 (1943), pp. 212–229.
Woodger, J. H. [Principles] Biological principles. New York, 1929.
Zilsel, Edgar. [Empiricism] Problems of empiricism. In *International encyclopedia of unified science*, vol. II, no. 8. The University of Chicago Press, 1941.
———. [Laws] Physics and the problem of historico-sociological laws. *Philosophy of science*, vol. 8 (1941), pp. 567–579.