Carl G. Hempel

ASPECTS of SCIENTIFIC EXPLANATION

And Other Essays in the Philosophy of Science

THE FREE PRESS, NEW YORK

COLLIER-MACMILLAN LIMITED, LONDON

ASPECTS OF

SCIENTIFIC EXPLANATION

CONTENTS

1. INTRODUCTION 333

2. DEDUCTIVE-NOMOLOGICAL EXPLANATION 335

2.1 Fundamentals: D-N Explanation and the Concept of Law 335

2.2 Causal Explanation and the D-N Model 347

2.3 The Role of Laws in Explanation 354

2.3.1. *The conception of laws as inference rules* 354

2.3.2. *The conception of laws as role-justifying grounds for explanations* 359

2.4 Explanation as Potentially Predictive 364

3. STATISTICAL EXPLANATION 376

3.1 Laws of Statistical Form 376

3.2 Deductive-Statistical Explanation 380

3.3 Inductive-Statistical Explanation 381

3.4 The Ambiguity of Inductive-Statistical Explanation and the Requirement of Maximal Specificity 394

3.4.1. *The problem of explanatory ambiguity* 394

3.4.2. *The requirement of maximal specificity and the epistemic relativity of inductive-statistical explanation* 397

3.4.3. *Discrete state systems and explanatory ambiguity* 403

3.5 Predictive Aspects of Statistical Explanation 406

3.6 The Nonconjunctiveness of Inductive-Statistical Explanation 410

4. THE CONCEPTS OF COVERING-LAW EXPLANATION AS EXPLICATORY MODELS 412
4.1 General Character and Intent of the Models 412
4.2 Varieties of Explanatory Incompleteness 415
4.2.1. *Elliptic formulation* 415
4.2.2. *Partial Explanation* 415
4.2.3. *Explanatory incompleteness vs. overdetermination* 418
4.2.4. *Explanatory incompleteness and "concrete events"* 421
4.2.5. *Explanatory closure: explanation sketch* 423
4.3 Concluding Remark on The Covering-Law Models 424

5. PRAGMATIC ASPECTS OF EXPLANATION 425
5.1 Introductory Remarks 425
5.2 Explaining How-Possibly 428
5.3 Explanation vs. Reduction to the Familiar 430

6. MODELS AND ANALOGIES IN SCIENTIFIC EXPLANATION 433

7. GENETIC EXPLANATION AND COVERING LAWS 447

8. EXPLANATION-BY-CONCEPT 453

9. DISPOSITIONAL EXPLANATION 457

10. THE CONCEPT OF RATIONALITY AND THE LOGIC OF EXPLANATION BY REASONS 463
10.1 Two Aspects of the Concept of Rationality 463
10.2 Rationality as a Normative-Critical Concept 463
10.3 Rationality as an Explanatory Concept 469
10.3.1. *Dray's concept of rational explanation* 469
10.3.2. *Explanation by reasons as broadly dispositional* 472
10.3.3. *Epistemic interdependence of belief attributions and goal attributions* 475
10.3.4. *Rational action as an explanatory model concept* 477
10.3.5. *The model of a consciously rational agent* 478
10.3.6. *The "rationality" of nondeliberative actions. Explanation by unconscious motives* 483
10.3.7. *A note on causal aspects of dispositional explanations* 486

11. CONCLUDING REMARKS 487

BIBLIOGRAPHY 489

1. INTRODUCTION

Among the many factors that have prompted and sustained inquiry in the diverse fields of empirical science, two enduring human concerns have provided the principal stimulus for man's scientific efforts.

One of them is of a practical nature. Man wants not only to survive in the world, but also to improve his strategic position in it. This makes it important for him to find reliable ways of foreseeing changes in his environment and, if possible, controlling them to his advantage. The formulation of laws and theories that permit the prediction of future occurrences are among the proudest achievements of empirical science; and the extent to which they answer man's quest for foresight and control is indicated by the vast scope of their practical applications, which range from astronomic predictions to meteorological, demographic, and economic forecasts, and from physico-chemical and biological technology to psychological and social control.

The second basic motive for man's scientific quest is independent of such practical concerns; it lies in his sheer intellectual curiosity, in his deep and persistent desire to know and to understand himself and his world. So strong, indeed, is this urge that in the absence of more reliable knowledge, myths are often invoked to fill the gap. But in time, many such myths give way to scientific conceptions of the what and the why of empirical phenomena.

What is the nature of the explanations empirical science can provide? What understanding of empirical phenomena do they convey? This essay attempts to shed light on these questions by examining in some detail the form and the function of some of the major types of explanatory account that have been advanced in different areas of empirical science.

The terms 'empirical science' and 'scientific explanation' will here be understood to refer to the entire field of empirical inquiry, including the natural and the social sciences as well as historical research. This broad use of the two terms is not intended to prejudge the question of the logical and methodological

This essay has not previously appeared in print. It includes, however, some passages from the following articles:

"Deductive-Nomological vs. Statistical Explanation," *Minnesota Studies in the Philosophy of Science*, Vol. III, edited by Herbert Feigl and Grover Maxwell. University of Minnesota Press, Minneapolis. Copyright 1962 by the University of Minnesota.—Excerpts reprinted by permission of the publisher.

"Explanation in Science and in History," R. Colodny (ed.) *Frontiers of Science and Philosophy*, Pittsburgh: University of Pittsburgh Press, 1962; pp. 9-33. Excerpts reprinted by permission of the publisher.

"Rational Action," from *Proceedings and Addresses of the American Philosophical Association*, Vol. 35 (1961-62), pp. 5-23. Yellow Springs, Ohio: The Antioch Press, 1962. Excerpts reprinted by permission of the American Philosophical Association.

similarities and differences between different areas of empirical inquiry, except for indicating that the procedures used in those different areas will be taken to conform to certain basic standards of objectivity. According to these standards, hypotheses and theories—including those invoked for explanatory purposes—must be capable of test by reference to publicly ascertainable evidence, and their acceptance is always subject to the proviso that they may have to be abandoned if adverse evidence or more adequate hypotheses or theories should be found.

A scientific explanation may be regarded as an answer to a why-question, such as: 'Why do the planets move in elliptical orbits with the sun at one focus?', 'Why does the moon look much larger when it is near the horizon than when it is high in the sky?', 'Why did the television apparatus on Ranger VI fail?', 'Why are children of blue-eyed parents always blue-eyed?', 'Why did Hitler go to war against Russia?'. There are other modes of formulating what we will call *explanation-seeking questions:* we might ask what caused the failure of the television apparatus on Ranger VI, or what led Hitler to his fateful decision. But a why-question always provides an adequate, if perhaps sometimes awkward, standard phrasing.

Sometimes the subject matter of an explanation, or the *explanandum*, is indicated by a noun, as when we ask for an explanation of the aurora borealis. It is important to realize that this kind of phrasing has a clear meaning only in so far as it can be restated in terms of why-questions. Thus, in the context of an explanation, the aurora borealis must be taken to be characterized by certain distinctive general features, each of them describable by a that-clause, for example: that it is normally found only in fairly high northern latitudes; that it occurs intermittently; that sunspot maxima, with their eleven-year cycle, are regularly accompanied by maxima in the frequency and brightness of aurora borealis displays; that an aurora shows characteristic spectral lines of rare atmospheric gases, and so on. And to ask for an explanation of the aurora borealis is to request an explanation of *why* auroral displays occur in the fashion indicated and *why* they have physical characteristics such as those just mentioned. Indeed, requests for an explanation of the aurora borealis, of the tides, of solar eclipses in general or of some individual solar eclipse in particular, or of a given influenza epidemic, and the like have a clear meaning only if it is understood what aspects of the phenomenon in question are to be explained; and in that case the explanatory problem can again be expressed in the form 'Why is it the case that *p*?', where the place of '*p*' is occupied by an empirical statement specifying the explanandum. Questions of this type will be called *explanation-seeking why-questions*.

Not all why-questions call for explanations, however. Some of them solicit reasons in support of an assertion. Thus, statements such as 'Hurricane Delila will veer out into the Atlantic', 'He must have died of a heart attack', 'Plato would have disliked Stravinsky's music' might be met with the question 'Why should this be so?', which seeks to elicit, not an explanation, but evidence or grounds or reasons in support of the given assertion. Questions of this kind will be called *reason-seeking* or *epistemic*. To put them into the form 'Why should it be the case that *p*?' is misleading; their intent is more adequately conveyed by a phrasing such as 'Why should it be believed that *p*?' or 'What reasons are there for believing that *p*?'.

An explanation-seeking why-question normally presupposes that the statement occupying the place of '*p*' is true, and asks for an explanation of the presumptive fact, event, or state of affairs described by it; an epistemic why-question does not presuppose the truth of the corresponding statement, but on the contrary, solicits reasons for believing it true. An appropriate answer to the former will therefore offer an explanation of a presumptive empirical phenomenon; whereas an appropriate answer to the latter will offer validating or justifying grounds in support of a statement. Despite these differences in presuppositions and objectives, there are also important connections between the two kinds of question; in particular, as will be argued later (in sections 2.4 and 3.5), any adequate answer to an explanation-seeking question 'Why is it the case that *p*?' must also provide a potential answer to the corresponding epistemic question 'What grounds are there for believing that *p*?'

In the discussion that follows, I will first distinguish two basic types of scientific explanation, deductive-nomological and inductive-statistical, each characterized by a schematic "model"; and I will examine certain logical and methodological questions to which these models give rise, including a number of objections that have been raised against them. Following this, I propose to assess the significance and adequacy of the basic conceptions inherent in those models by exploring the extent to which they can serve to analyze the structure and to illuminate the rationale of different kinds of explanation offered in empirical science.

2. DEDUCTIVE-NOMOLOGICAL EXPLANATION

21. FUNDAMENTALS: D-N EXPLANATION AND THE CONCEPT OF LAW. In his book, *How We Think*,[1] John Dewey describes a phenomenon he observed one day while washing dishes. Having removed some glass tumblers from the hot suds and placed them upside down on a plate, he noticed that soap bubbles emerged from under the tumbler's rims, grew for a while, came to a standstill

1. Dewey (1910), chap. VI.

and finally receded into the tumblers. Why did this happen? Dewey outlines an explanation to this effect: Transferring the tumblers to the plate, he had trapped cool air in them; that air was gradually warmed by the glass, which initially had the temperature of the hot suds. This led to an increase in the volume of the trapped air, and thus to an expansion of the soap film that had formed between the plate and the tumblers' rims. But gradually, the glass cooled off, and so did the air inside, and as a result, the soap bubbles receded.

The explanation here outlined may be regarded as an argument to the effect that the phenomenon to be explained, *the explanandum phenomenon*, was to be expected in virtue of certain explanatory facts. These fall into two groups: (i) particular facts and (ii) uniformities expressible by means of general laws. The first group includes facts such as these: the tumblers had been immersed in soap suds of a temperature considerably higher than that of the surrounding air; they were put, upside down, on a plate on which a puddle of soapy water had formed that provided a connecting soap film, and so on. The second group of explanatory facts would be expressed by the gas laws and by various other laws concerning the exchange of heat between bodies of different temperature, the elastic behavior of soap bubbles, and so on. While some of these laws are only hinted at by such phrasings as 'the warming of the trapped air led to an increase in its pressure', and others are not referred to even in this oblique fashion, they are clearly presupposed in the claim that certain stages in the process yielded others as their results. If we imagine the various explicit or tacit explanatory assumptions to be fully stated, then the explanation may be conceived as a deductive argument of the form

$$\text{(D-N)} \qquad \begin{array}{l} \left.\begin{array}{c} C, C_2, \ldots, C_k \\ \\ L_1, L_2, \ldots, L_r \end{array}\right\} \text{Explanans } S \\ \hline \qquad\qquad E \qquad\qquad \text{Explanandum-sentence} \end{array}$$

Here, $C_1, C_2, \ldots, C_k$ are sentences describing the particular facts invoked; $L_1, L_2, \ldots, L_r$ are the general laws on which the explanation rests. Jointly these sentences will be said to form the *explanans* S, where S may be thought of alternatively as the set of the explanatory sentences or as their conjunction. The conclusion E of the argument is a sentence describing the explanandum-phenomenon; I will call E the explanandum-sentence or explanandum-statement; the word 'explanandum' alone will be used to refer either to the explanandum-phenomenon or to the explanandum-sentence: the context will show which is meant.

The kind of explanation whose logical structure is suggested by the schema

(D-N) will be called *deductive-nomological explanation* or D-N *explanation* for short; for it effects a deductive subsumption of the explanandum under principles that have the character of general laws. Thus a D-N explanation answers the question '*Why* did the explanandum-phenomenon occur?' by showing that the phenomenon resulted from certain particular circumstances, specified in $C_1, C_2, \ldots, C_k$, in accordance with the laws $L_1, L_2, \ldots, L_r$. By pointing this out, the argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected;* and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred.[2]

In a D-N explanation, then, the explanandum is a logical consequence of the explanans. Furthermore, reliance on general laws is essential to a D-N explanation; it is in virtue of such laws that the particular facts cited in the explanans possess explanatory relevance to the explanandum phenomenon. Thus, in the case of Dewey's soap bubbles, the gradual warming of the cool air trapped under the hot tumblers would constitute a mere accidental antecedent rather than an explanatory factor for the growth of the bubbles, if it were not for the gas laws, which connect the two events. But what if the explanandum sentence E in an argument of the form (D-N) is a logical consequence of the sentences $C_1, C_2, \ldots, C_k$ alone? Then, surely, no empirical laws are *required* to deduce E from the explanans; and any laws included in the latter are gratuitous, dispensable premises. Quite so; but in this case, the argument would not count as an explanation. For example, the argument:

$$\frac{\text{The soap bubbles first expanded and then receded}}{\text{The soap bubbles first expanded}}$$

2. A general conception of scientific explanation as involving a deductive subsumption under general laws was espoused, though not always clearly stated, by various thinkers in the past, and has been advocated by several recent or contemporary writers, among them N. R. Campbell [(1920), (1921)], who developed the idea in considerable detail. In a textbook published in 1934, the conception was concisely stated as follows: "Scientific explanation consists in subsuming under some rule or law which expresses an invariant character of a group of events, the particular events it is said to explain. Laws themselves may be explained, and in the same manner, by showing that they are consequences of more comprehensive theories." (Cohen and Nagel 1934, p. 397.) Popper has set forth this construal of explanation in several of his publications; *cf.* the note at the end of section 3 in Hempel and Oppenheim (1948). His earliest statement appears in section 12 of his book (1935), of which his work (1959) is an expanded English version. His book (1962) contains further observations on scientific explanation. For some additional references to other proponents of the general idea, see Donagan (1957), footnote 2; Scriven (1959), footnote 3. However, as will be shown in section 3, deductive subsumption under general laws does not constitute the only form of scientific explanation.

though deductively valid, clearly cannot qualify as an explanation of why the bubbles first expanded. The same remark applies to all other cases of this kind. A D-N explanation will have to contain, in its explanans, some general laws that are *required* for the deduction of the explanandum, i.e. whose deletion would make the argument invalid.

If the explanans of a given D-N explanation is true, i.e. if the conjunction of its constituent sentences is true, we will call the *explanation true;* a true explanation, of course, has a true explanandum as well. Next, let us call a *D-N explanation more or less strongly supported or confirmed* by a given body of evidence according as its explanans is more or less strongly confirmed by the given evidence. (One factor to be considered in appraising the empirical soundness of a given explanation will be the extent to which its explanans is supported by the total relevant evidence available.) Finally, by a *potential D-N explanation*, let us understand any argument that has the character of a D-N explanation except that the sentences constituting its explanans need not be true. In a potential D-N explanation, therefore, $L_1, L_2, \ldots, L_r$ will be what Goodman has called *lawlike sentences*, i.e. sentences that are like laws except for possibly being false. Sentences of this kind will also be referred to as *nomic* or *nomological*. We use the notion of a potential explanation, for example, when we ask whether a novel and as yet untested law or theory would provide an explanation for some empirical phenomenon; or when we say that the phlogiston theory, though now discarded, afforded an explanation for certain aspects of combustion.[3] Strictly speaking, only true lawlike statements can count as laws—one would hardly want to speak of false laws of nature. But for convenience I will occasionally use the term 'law' without implying that the sentence in question is true, as in fact, I have done already in the preceding sentence.

The characterization of laws as true lawlike sentences raises the important and intriguing problem of giving a clear characterization of lawlike sentences without, in turn, using the concept of law. This problem has proved to be highly recalcitrant, and I will make here only a few observations on certain aspects of it that are relevant also to the analysis of scientific explanation.

Lawlike sentences can have many different logical forms. Some paradigms of nomic sentences, such as 'All gases expand when heated under constant pressure' may be construed as having the simple universal conditional form '$(x)(Fx \supset Gx)$'; others involve universal as well as existential generalization,

3. The explanatory role of the phlogiston theory is described in Conant (1951), pp. 164-71. The concept of potential explanation was introduced in Hempel and Oppenheim (1948), section 7. The concept of lawlike sentence, in the sense here indicated, is due to Goodman (1947).

as does the sentence 'For every chemical compound there exists a range of temperatures and pressures at which the compound is liquid'; many of the lawlike sentences and theoretical principles of the physical sciences assert more or less complex mathematical relationships between different quantitative variables.[4]

But lawlike sentences cannot be characterized in terms of their form alone. For example, not all sentences of the simple universal conditional form just mentioned are lawlike; hence, even if true, they are not laws. The sentences 'All members of the Greenbury School Board for 1964 are bald' and 'All pears in this basket are sweet' illustrate this point. Goodman[5] has pointed out a characteristic that distinguishes laws from such nonlaws: The former can, whereas the latter cannot, sustain counterfactual and subjunctive conditional statements. Thus the law about the expansion of gases can serve to support statements such as 'If the oxygen in this cylinder had been heated (were heated) under constant pressure then it would have expanded (would expand)'; whereas the statement about the School Board lends no support at all to the subjunctive conditional 'If Robert Crocker were a member of the Greenbury School Board for 1964 then he would be bald'.

We might add that the two kinds of sentence differ analogously in explanatory power. The gas law, in combination with suitable particular data, such as that the oxygen in the cylinder was heated under constant pressure, can serve to explain why the volume of the gas increased; but the statement about the School Board, analogously combined with a statement such as 'Harry Smith is a member of the Greenbury School Board for 1964' cannot explain why Harry Smith is bald.

But though these observations shed light on the concept of lawlikeness they afford no satisfactory explication of it; for one of them presupposes an understanding of counterfactual and of subjunctive conditional statements, which present notorious philosophical difficulties; the other makes use of the idea of explanation to clarify the concept of a lawlike statement; and we are

4. Fain (1963), p. 524, strangely claims that "Hempel and Oppenheim failed to consider" (in their essay, 1948) "generalizations that are basically of the form $(x)\ (\exists y)\ Pxy$". But in section 7 of the essay in question, we specifically admitted laws and theories of any of the quantificational types expressible in the lower functional calculus, and we required that they be essentially generalized sentences containing "one or more quantifiers." Similarly, when Scriven speaks of "the deductive model, with its syllogistic form, where no student of elementary logic could fail to complete the inference, given the premise" (1959, p. 462), he imposes upon the model an entirely unwarranted oversimplified construal; for the schema (D-N) clearly allows for the use of highly complex general laws of the kind specified in the text above; and where these occur in the explanans, the explanandum cannot, of course, be deduced by syllogistic methods.

5. Goodman (1955), p. 25; for certain qualifications, *cf. ibid.*, p. 118.

here trying conversely to characterize a certain type of explanation with the help of concepts which include that of lawlike statement.

Now, our examples of non-lawlike sentences share a characteristic that might seem to afford a criterion for the distinction we seek to draw; namely, each of them applies to only a finite number of individual cases or instances. Must not a general law be conceived as admitting of indefinitely many instances?

Surely a lawlike sentence must not be *logically* limited to a finite number of instances: it must not be logically equivalent to a finite conjunction of singular sentences, or, briefly, it must be of *essentially generalized form*. Thus, the sentence 'Every element of the class consisting of the objects a, b, and c has the property P' is not lawlike; for it is logically equivalent to the conjunction '$Pa \cdot Pb \cdot Pc$', and clearly a sentence of this kind cannot support counterfactual conditionals or provide explanations.[6]

But our two earlier nonlawlike generalizations are not ruled out by this condition: they are not logically equivalent to corresponding finite conjunctions since they do not state specifically who are the members of the School Board, or what particular pears are in the basket. Should we, then, deny lawlike status also to any general sentence which—by empirical accident, so to speak—has only a finite number of instances? This would surely be ill-advised. Suppose, for example, that from the basic laws of celestial mechanics a general statement is derived concerning the relative motion of the components of a double star in the special case where those components are of exactly equal mass. Is this statement to be termed a law only if it has been established that there exist at least two (or perhaps more) instances of this special kind of double star? Or consider the general statement, derivable from Newton's laws of gravitation and of motion, which deals, in a manner similar to Galileo's law, with the free fall of physical bodies near the surface of a spherical mass having the same density as the Earth, but twice its radius. Should this statement not be called a law unless it had been shown to have several instances—even though it is a logical consequence of a set of laws with many instances?

6. In such references to "the form" of a sentence, there lurks another difficulty: that form is clearly determined only if the sentence is expressed in a formalized language. An English sentence such as 'This object is soluble in water' may be construed as a singular sentence of the form 'Pa', but alternatively also as a sentence of generalized form stating that if at any time the object is put into any (sufficiently large) body of water, it will dissolve. (This will be elaborated further in section 2.3.1.) Our remark about a sentence of the form 'For all x, if x is a, b, or c, then x has property P' might be stated more circumspectly by saying that that kind of sentence is not a law *in terms of P*; it cannot serve to explain the occurrence of P in any particular case; nor can it support counterfactual or subjunctive conditionals about particular occurrences of P.

Besides, there appears to be only an inessential "difference in degree" between a general statement that happens to have just one instance and another which happens to have two or some other finite number. But, then, how many instances would a law be required to have? To insist on some particular finite number would be arbitrary; and the requirement of an infinite number of actual instances would raise obvious difficulties. Clearly, the concept of scientific law cannot reasonably be subjected to any condition concerning the number of instances, except for the requirement barring logical equivalence with singular statements.

Besides, we should note that the concept, presupposed in the preceding discussion, of a "case" or an "instance" of a general statement is by no means as clear as it might seem. Consider, for example, general statements of the form, 'All objects with the property *F* also have the property *G*', or briefly 'All *F* are *G*'. It seems natural to accept the criterion that a particular object *i* is an instance of such a statement if and only if *i* has the property *F* and the property *G*, or briefly, if *i* is both *F* and *G*. This would imply that if there are no objects with the property *F* at all, the general statement has no instances. Yet, the statement is logically equivalent with 'All non-*G* are non-*F*', which, under the contemplated criterion, may well have instances even if there are no *F*. Thus, the general statement, 'All unicorns feed on clover' would have no instance, but its equivalent 'Anything that does not feed on clover is not a unicorn' would have many—perhaps infinitely many—instances. An analogous remark might well be true of the law mentioned earlier concerning double stars whose components have equal mass. Hence, the contemplated criterion of instantiation, which seems quite obvious at first, has the consequence that of two logically equivalent general statements, one may have no instances, the other, infinitely many. But this makes the criterion unacceptable since such equivalent sentences express the same law and thus should be instantiated by the same objects.

For laws of the simple kind just considered, the following alternative definition of instantiation will suffice to assign the same instances to equivalent statements: an object *i* is an instance of the statement 'All *F* are *G*' if and only if it is not the case that *i* is *F* but not *G*. However, for laws of more complex logical form, the concept of instance raises further problems.[7] But these

7. These difficulties concerning the intuitive idea of instantiation of a general law are closely related to the paradoxes of confirmation set forth in Hempel (1945). The inadequacy of the initially contemplated intuitive criterion is further illustrated by the following consequence: The sentence 'All *F* are *G*' is logically equivalent to 'Anything that is *F* but not *G* is both *G* and not *G*'; and on the criterion in question, this sentence clearly cannot have any instances—even if 'All *F* are *G*' is true and is instantiated by infinitely many objects that are both *F* and *G*. Our modified criterion of instantiation avoids this difficulty: the sets of instances, thus construed, of any two logically equivalent universally quantified sentences in one variable are identical.

need not be pursued here, for I am not proposing that a law must satisfy certain minimum conditions concerning the number of its instances.

There is yet another common trait of our non-lawlike generalizations that seems to hold promise as a criterion for the distinction here under discussion: they contain terms, such as 'this basket' and 'the Greenbury School Board for 1964', which directly or indirectly refer to particular objects, persons, or places; whereas the terms occurring in Newton's laws or in the gas laws involve no such reference. In an earlier article on the subject, Oppenheim and I suggested, therefore, that the constituent predicates of what we called fundamental lawlike sentences must all be such that the specification of their meaning requires no reference to any one particular object or location.[8] We noted, however, that this characterization still is not satisfactory for purposes of explication because the idea of "the meaning" of a given term is itself far from being clear.

Besides, reference to particular individuals does not always deprive a general statement of explanatory power, as is illustrated by Galileo's law for free fall, whose full formulation makes reference to the earth. Now it is true that, with qualifications soon to be stated, Galileo's law may be regarded as derivable from the laws of Newtonian theory, which have the character of fundamental lawlike sentences, so that an explanation based on Galileo's law can also be effected by means of fundamental laws. But it certainly cannot be taken for granted that all other laws mentioning particular individuals can similarly be derived from fundamental laws.

Goodman, in a searching exploration of the concept of law, has argued that, in contrast to non-lawlike generalizations, lawlike sentences are capable of being supported by observed instances and hence of being "projected" from examined to unexamined cases; and he has argued further that the relative "projectibiltiy" of generalizations is determined primarily by the relative "entrenchment" of their constituent predicates, i.e. by the extent to which those predicates have been used in previously projected generalizations.[9] Thus, terms, like 'member of the Greenbury School Board for 1964' and 'pear in this basket' would be disqualified, for the purposes of formulating lawlike sentences, on the ground that they lack adequate entrenchment.

8. Hempel and Oppenheim (1948), section 6. "Specification of meaning" might be conceived as effected by definition or perhaps by weaker means, such as Carnap's reduction sentences. See Carnap (1938) and, for more details, (1936-37). The distinction thus attempted between those terms which in some way refer to particular individuals and those which do not is closely akin to the distinction made by Popper, in section 14 of (1935) and (1959), between individual concepts, "in the definition of which proper names (or equivalent signs) are indispensable," and universal concepts, for which this is not the case.

9. For details, and for further considerations that affect projectibility, see Goodman (1955), especially chapters III and IV.

But while Goodman's criterion thus succeeds in barring from the class of lawlike sentences such generalizations as our two examples, the class of lawlike sentences it delimits still seems too inclusive for our purposes. For according to Goodman, the "entrenchment of a predicate results from the actual projection not merely of that predicate alone but also of all predicates coextensive with it. In a sense, not the word itself but the class it selects is what becomes entrenched . . ."[10] Hence, replacing a predicate in a lawlike sentence by a coextensive one should yield a lawlike sentence again. Is this generally the case? Suppose that the hypothesis h: '$(x)(Px \supset Qx)$' is lawlike, but that as a matter of empirical fact there happen to be just three elements in the class selected by 'P', namely a, b, and c. Then 'Px' is coextensive with '$x = a \vee x = b \vee x = c$.' Replacement of '$Px$' by this expression, however, turns h into the sentence '$(x)[(x=a \vee x=b \vee x=c) \supset Qx]$', which, being logically equivalent with '$Qa \cdot Qb \cdot Qc$', is not lawlike on our understanding that a lawlike sentence must be of essentially generalized form, so as to be able to serve in an explanatory role. Our conception of lawlikeness differs at this point from that envisaged by Goodman, who introduces the notion principally in an effort to establish a dividing line between sentences that are confirmable by their instances and those that are not.[11] It may not be necessary to require of the former that they be of essentially general form, and Goodman does not impose this requirement on lawlike sentences. For laws, however, that are to function in an explanatory capacity, the requirement seems to me indispensable.

Though the preceding discussion has not led to a fully satisfactory general characterization of lawlike sentences and thus of laws, it will, I hope, have clarified to some extent the sense in which those concepts will be understood in the present study.[12]

The examples we have considered so far illustrate the deductive explanation of particular occurrences by means of empirical laws. But empirical science raises the question "Why?" also in regard to the uniformities expressed by such laws and often answers it, again, by means of a deductive-nomological explanation, in which the uniformity in question is subsumed under more inclusive laws or under theoretical principles. For example, the questions of why freely falling bodies move in accordance with Galileo's law and why the motion of the planets exhibit the uniformities expressed by Kepler's laws are answered by showing that these laws are but special consequences of the Newtonian laws of gravitation and of motion. Similarly, the uniformities

10. Goodman (1955), pp. 95-96.

11. On this distinction, see the Postscript to the article "Studies in the Logic of Confirmation" in this volume.

12. For further discussions of the problems here referred to see Braithwaite (1953), chap. IX and Nagel (1961), chap. 4.

expressed by the laws of geometrical optics, such as those of the rectilinear propagation of light and of reflection and refraction, are accounted for by subsumption under the principles of wave optics. For brevity, an explanation of a uniformity expressed by a law will sometimes be elliptically referred to as an explantion of the law in question.

It should be noted, however, that in the illustrations just mentioned, the theory invoked does not, strictly speaking, imply the presumptive general laws to be explained; rather, it implies that those laws hold only within a limited range, and even there, only approximately. Thus, Newton's law of gravitation implies that the acceleration of a freely falling body is not constant, as Galileo's law asserts, but undergoes a very slight but steady increase as the body approaches the ground. But while, strictly speaking, Newton's law contradicts Galileo's, it shows that the latter is almost exactly satisfied in free fall over short distances. In slightly greater detail, we might say that the Newtonian theory of gravitation and of motion implies its own laws concerning free fall under various circumstances. According to one of these, the acceleration of a small object falling freely toward a homogeneous spherical body varies inversely as the square of its distance from the center of the sphere, and thus increases in the course of the fall; and the uniformity expressed by this law is explained in a strictly deductive sense by the Newtonian theory. But when conjoined with the assumption that the earth is a homogeneous sphere of specified mass and radius, the law in question implies that for free fall over short distances near the surface of the earth, Galileo's law holds to a high degree of approximation; in this sense, the theory might be said to provide an *approximative D-N explanation* of Galileo's law.

Again, in the case of planetary motion, the Newtonian theory implies that since a planet is subject to gravitational attraction not only from the Sun, but also from the other planets, its orbit will not be exactly elliptical, but will show certain perturbations. Hence, as Duhem[13] noted, Newton's law of gravitation, far from being an inductive generalization based on Kepler's laws, is, strictly speaking, incompatible with them. One of its important credentials is precisely the fact that it enables the astronomer to compute the deviations of the planets from the elliptic orbits Kepler had assigned to them.

A similar relation obtains between the principles of wave optics and the laws of geometrical optics. For example, the former calls for a diffractive "bending" of light around obstacles—a phenomenon ruled out by the con-

13. See Duhem (1906), pp. 312 ff. Duhem's remarks on this subject are included in those excerpts from P. P. Wiener's translation of Duhem's work that are reprinted in Feigl and Brodbeck (1953). The point has recently been re-emphasized by several writers, among them Popper (1957a), pp. 29-34, and Feyerabend (1962), pp. 46-48.

ception of light as composed of rays traveling in straight lines. But in analogy to the preceding illustration, the wave-theoretical account implies that the laws of rectilinear propagation, of reflection, and of refraction as formulated in geometrical optics are satisfied to a very high degree of approximation within a limited range of cases, including those which provided experimental support for the laws in their original formulation.

In general, an explanation based on theoretical principles will both broaden and deepen our understanding of the empirical phenomena concerned. It will achieve an increase in breadth because the theory will usually cover a wider range of occurrences than do the empirical laws previously established. For example, Newton's theory of gravitation and of motion governs free fall not only on the earth, but also on other celestial bodies; and not only planetary motions, but also the relative motion of double stars, the orbits of comets and of artificial satellites, the movements of pendulums, certain aspects of the tides, and many other phenomena. And a theoretical explanation deepens our understanding for at least two reasons. First, it reveals the different regularities exhibited by a variety of phenomena, such as those just mentioned in reference to Newton's theory, as manifestations of a few basic laws. Secondly, as we noted, the generalizations previously accepted as correct statements of empirical regularities will usually appear as approximations only of certain lawlike statements implied by the explanatory theory, and to be very nearly satisfied only within a certain limited range. And in so far as tests of the laws in their earlier formulation were confined to cases in that range, the theoretical account also indicates why those laws, though not generally true, should have been found confirmed.

When a scientific theory is superseded by another in the sense in which classical mechanics and electrodynamics were superseded by the special theory of relativity, then the succeeding theory will generally have a wider explanatory range, including phenomena the earlier theory could not account for; and it will as a rule provide approximative explanations for the empirical laws implied by its predecessor. Thus, special relativity theory implies that the laws of the classical theory are very nearly satisfied in cases involving motion only at velocities which are small compared to that of light.

The general conception of explanation by deductive subsumption under general laws or theoretical principles, as it has been outlined in this section, will be called the *deductive nomological-model*, or the *D-N model of explanation;* the laws invoked in such an explanation will also be referred to, in William Dray's suggestive phrase, as *covering laws*.[14] Unlike Dray, however, I will not

14. For Dray's use of the terms 'covering law' and 'covering law model', see Dray (1957), and also (1963), p. 106.

refer to the D-N model as the covering-law model, for I will subsequently introduce a second basic model of scientific explanation which also relies on covering laws, but which is not of deductive-nomological form. The term 'covering-law model' will then serve to refer to both of those models.

As the schema (D-N) plainly indicates, a deductive-nomological explanation is not conceived as invoking only one covering law; and our illustrations show how indeed many different laws may be invoked in explaining one phenomenon. A purely logical point should be noted here, however. If an explanation is of the form (D-N), then the laws $L_1, L_2, \ldots, L_r$ invoked in its explanans logically imply a law L^* which by itself would suffice to explain the explanandum event by reference to the particular conditions noted in the sentences $C_1, C_2, \ldots, C_k$. This law L^* is to the effect that whenever conditions of the kind described in the sentences $C_1, C_2, \ldots, C_k$ are realized then an event of the kind described by the explanandum-sentence occurs.[15] Consider an example: A chunk of ice floats in a large beaker of water at room temperature. Since the ice extends above the surface, one might expect the water level to rise as the ice melts; actually, it remains unchanged. Briefly, this can be explained as follows: According to Archimedes' principle, a solid body floating in a liquid displaces a volume of liquid that has the same weight as the body itself. Hence, the chunk of ice has the same weight as the water displaced by its submerged portion. Since melting does not change the weight, the ice turns into a mass of water of the same weight, and hence also of the same volume, as the water initially displaced by its submerged portion; consequently, the water level remains unchanged. The laws on which this account is based include Archimedes' principle, a law concerning the melting of ice at room temperature; the principle of the conservation of mass; and so on. None of these laws mentions the particular glass of water or the particular piece of ice with which the explanation is concerned. Hence the laws imply not only that as this particular piece of ice melts in this particular glass, the water level remains unchanged, but rather the general statement L^* that under the same *kind* of circumstance, i.e., when any piece of ice floats in water in any glass at room temperature, the same *kind* of phenomenon will occur, i.e., the water level will remain unchanged. The law L^* will usually be "weaker" than the laws $L_1, L_2, \ldots, L_r$; i.e., while being logically implied by the conjunction of those laws, it will not, in general, imply that conjunction. Thus, in our illustration one of the original explanatory laws applies also to the floating of a piece of marble on mercury or of a boat on water, whereas L^* deals only with the case of ice floating on water. But clearly, L^* in conjunction with $C_1, C_2, \ldots, C_k$ logically implies E and could indeed be used to

15. This was noted already in Hempel (1942), section 2.1.

explain, in this context, the event described by E. We might therefore refer to L^* as a *minimal covering law* implicit in a given D-N explanation.[16] But while such laws might be used for explanatory purposes, the D-N model by no means restricts deductive-nomological explanations to the use of minimal laws. Indeed such a restriction would fail to do justice to one important objective of scientific inquiry, namely, that of establishing laws and theories of broad scope, under which narrower generalizations may then be subsumed as special cases or as close approximations of such.[17]

2.2 Causal Explanation and the D-N Model. An explanation of a particular occurrence is often conceived as pointing out what "caused" it. Thus, the initial expansion of the soap bubbles described by John Dewey might be said to have been caused by the warming of the air caught in the tumblers. But causal attributions of this kind presuppose appropriate laws, such as that under

16. The problem of formulating a precise definition of this notion need not detain us: it can be solved only by reference to some formalized language, and for our purposes the rough characterization here given will suffice. Incidentally, the notion of "the number of laws" invoked in a given explanation is not as clear as it might seem, for one law may sometimes be quite plausibly rewritten as a conjunction of two or more, and, conversely, several laws may sometimes be plausibly conjoined into one. But again, it is not necessary for us to pursue this problem.

17. In a recent essay, Feyerabend has criticized the deductive model of explanation for leading "to the demand. . . that all successful theories in a given domain must be mutually consistent" (1962, p. 30), or, more fully, that "only such theories are admissible (for explanation and prediction) in a given domain which either *contain* the theories already used in this domain, or are at least *consistent* with them" (1962, p. 44, italics the author's). Feyerabend rightly argues that this demand conflicts with actual scientific procedure and is unsound on methodological grounds. But he is completely mistaken in his allegation—for which he offers no support—that the conception of explanation by deductive subsumption under general laws or theoretical principles entails the incriminated methodological maxim. Indeed, the D-N model of explanation concerns simply the relation between explanans and explanandum and implies nothing whatever about the compatibility of different explanatory principles that might be accepted successively in a given field of empirical science. In particular, it does not imply that a new explanatory theory may be accepted only on condition that it be logically compatible with those previously accepted. One and the same phenomenon, or set of phenomena, may be deductively subsumable under different, and logically incompatible, laws or theories. To illustrate this schematically: the fact that three objects a, b, c, each of which has the property P, also have the property Q could be deductively accounted for by the hypothesis H_1 that all and only P's are Q's, and alternatively by the hypothesis H_2 that all P's and also some non-P's are Q's; i.e., the explanandum-sentence '$Qa.Qb.Qc$' can be deduced from '$Pa.Pb.Pc$' in conjunction with either H_1 or H_2, although H_1 and H_2 are logically incompatible. Thus a "new" explanatory theory for a given class of phenomena may deductively account for those phenomena even though it is logically incompatible with an earlier theory which also deductively accounts for them. But the conflicting theories cannot both be true, and it may well be that the earlier theory is false. Hence the maxim criticized by Feyerabend is indeed unsound. But this observation does not affect the D-N model of explanation, which does not imply that maxim at all.

constant pressure the volume of a gas increases as its temperature rises. And by virtue of thus presupposing general laws which connect "cause" and "effect," causal explanation conforms to the D-N model. Let me briefly amplify and substantiate this remark.

Consider first the explanatory use of what may be called *general* statements of causal connection: these are to the effect that an event of some *kind* A (e.g., motion of a magnet through a closed wire loop) causes an event of a certain other *kind*, B (e.g., flow of an electric current in the wire). Without entering into a more detailed analysis, we may say that in the simplest case a statement of this type affirms a law to the effect that whenever an event of kind A takes place then there occurs, at the same location or at a specifiable different one, a corresponding event of kind B. This construal fits, for example, the statements that motion of a magnet causes the flow of a current in a neighboring wire loop, and that raising the temperature of a gas under constant pressure increases its volume. Many general statements of causal connection call for a more complex analysis, however. Thus, the statement that in a mammal, stoppage of the heart will cause death presupposes certain "standard" conditions that are not explicitly stated, but that are surely meant to preclude, for example, the use of a heart-lung machine. "To say that X causes Y is to say that under proper conditions, an X will be followed by a Y," as Scriven[18] puts it. When this kind of causal locution is used, there usually is some understanding of what "proper" or "standard" background conditions are presupposed in the given context. But to the extent that those conditions remain indeterminate, a general statement of causal connection amounts at best to the vague claim that *there are* certain further unspecified background conditions whose explicit mention in the given statement would yield a truly general law connecting the "cause" and the "effect" in question.

Next, consider statements of causal connections between individual events. Take, for example, the assertion that the expansion and subsequent shrinkage of Dewey's soap bubbles were *caused* by a rise and subsequent drop of the temperature of the air trapped in the tumblers. Clearly, those temperature changes afford the requisite explanation only in conjunction with certain other conditions, such as the presence of a soap film, practically constant temperature and pressure of the air outside the glasses, and so on. Accordingly, in the context of explanation, a "cause" must be allowed to be a more or less complex set of circumstances and events, which might be described by a set of statements $C_1, C_2, \ldots, C_k$. And, as is suggested by the principle "Same cause, same effect," the assertion that those circumstances jointly caused a given event implies that

18. Scriven (1958), p. 185.

whenever and wherever circumstances of the kind in question occur, an event of the kind to be explained takes place. Thus the causal explanation implicitly claims that there are general laws—let us say, $L_1, L_2, \ldots, L_r$—in virtue of which the occurrence of the causal antecedents mentioned in $C_1, C_2, \ldots, C_k$ is a sufficient condition for the occurrence of the explanandum event. This relation between causal factors and effect is reflected in our schema (D-N): causal explanation is, at least implicitly, deductive-nomological.

Let me restate the point in more general terms. When an individual event *b* is said to have been caused by another individual event *a*, then surely the claim is implied that whenever "the same cause" is realized, "the same effect" will occur. But this claim cannot be taken to mean that whenever *a* recurs then so does *b;* for *a* and *b* are individual events at particular spatiotemporal locations and thus occur only once. Rather, *a* and *b* must be viewed as particular events of certain *kinds* (such as heating or cooling of a gas, expansion or shrinking of a gas) of which there may be further instances. And the law tacitly implied by the assertion that *b*, as an event of kind *B*, was caused by *a* as an event of kind *A* is a general statement of causal connection to the effect that, under suitable circumstances, an instance of *A* is invariably accompanied by an instance of *B*. In most causal explanations the requisite circumstances are not fully stated; the import of the claim that *b* was caused by *a* may then be suggested by the following approximate formulation: Event *b* was in fact preceded by event *a* in circumstances which, though not fully specified, were of such a kind that an occurrence of an event of kind *A* under such circumstances is universally followed by an event of kind *B*. For example, the statement that the burning (event of kind *B*) of a particular haystack was caused by a lighted cigarette dropped into the hay (particular event of kind *A*) asserts, first of all, that the latter event did take place; but a burning cigarette will set a haystack on fire only if certain further conditions are satisfied, which cannot at present be fully stated; and thus, the causal attribution at hand implies secondly that further conditions of a not fully specified kind were realized, under which an event of kind *A* is invariably followed by an event of kind *B*.

To the extent that a statement of individual causation leaves the relevant antecedent conditions, and thus also the requisite explanatory laws, indefinite it is like a note saying that there is a treasure hidden somewhere. Its significance and utility will increase as the location of the treasure is more narrowly circumscribed, as the relevant conditions and the corresponding covering laws are made increasingly explicit. In some cases, this can be done quite satisfactorily; the covering-law structure then emerges, and the statement of individual causal connection becomes amenable to test. When, on the other hand, the relevant conditions or laws remain largely indefinite, a statement of causal connection is rather in the nature of a program, or of a sketch, for an explanation in terms of causal laws; it might also be viewed as a "working hypothesis" which may prove its worth by giving new, and fruitful, direction to further research.

The view here taken of statements of individual causation might be further clarified by some comments on the thesis that "when one asserts that *X* causes *Y* one is certainly committed to the generalization that an identical cause would produce an identical effect, but this in no way commits one to any necessity for producing laws not involving the term 'identical,' which justify this claim. Producing laws is one way, not necessarily more conclusive, and usually less easy than other ways of supporting the causal statement. . . . (The idea of individual causation has, I think, this not inconsiderable basis.)"[19] Two questions must be clearly distinguished here, namely (i) what is being claimed by the statement that *X* causes *Y* (where, in the case of "individual causation," *X* and *Y* are individual events), and in particular, whether asserting it commits one to a generalization, and (ii) what kind of evidence would support the causal statement, and in particular, whether support can be provided only by producing generalizations in the form of laws.

Concerning the first question, I have argued that the given causal statement must be taken to claim by implication that an appropriate law or set of laws holds by virtue of which *X* causes *Y*. But, as noted earlier, the laws in question *cannot* be expressed by saying that an identical cause would produce an identical effect; for if *X* and *Y* are individual events with specific spatiotemporal locations, the recurrence of a cause identical with *X*, or of an effect identical with *Y*, is logically impossible. Rather, the general claim implied by the statement of individual causation that *X* caused *Y* is of the kind suggested in our discussion of the assertion that individual event *a*, as an instance of *A*, caused individual event *b*, as an instance of B.

We turn now to the second question. In certain cases, such as that of the soap bubbles observed by Dewey, some of the laws connecting the individual events *X* and *Y* may be explicitly stateable; and then, it may be possible to secure supporting evidence for them by appropriate experiments or observations. Hence, while the statement of individual causal connection implicitly *claims* the existence of underlying laws, the claim may well be *supported* by evidence consisting of particular confirming instances rather than of general laws. In other cases, when the nomological claim implicit in a causal statement is merely to the effect that *there are* relevant factors and suitable laws connecting *X* and *Y*, it may be possible to lend some credibility to this claim by showing

19. Scriven (1958), p. 194.

that under certain conditions, an event of kind X is at least very frequently accompanied by an event of kind Y: this might justify the working hypothesis that the background conditions could be further narrowed down in a way that would eventually yield a strictly causal connection. It is this kind of statistical evidence, for example, that is adduced in support of such claims as that cigarette smoking is "a cause of" or "a causative factor in" cancer of the lungs. In this case, the supposed causal laws cannot at present be explicitly stated. Thus, the nomological claim implied by this causal conjecture is of the existential type; it has the character of a working hypothesis for further research. The statistical evidence adduced lends support to the hypothesis and suggests further investigation, aimed at determining more precisely the conditions under which smoking will lead to cancer of the lungs.

The best examples of explanations conforming to the D-N model are based on physical theories of deterministic character. Briefly, a deterministic theory deals with the changes of "state" in physical systems of some specified kind. The state of such a system at any given time is characterized by the values assumed at that time by certain quantitative characteristics of the system, the so-called variables of state; and the laws specified by such a theory for the changes of state are deterministic in the sense that, given the state of the system at any one time, they determine its state at any other, earlier or later, time. For example, classical mechanics offers a deterministic theory for a system of point masses (or, practically, bodies that are small in relation to their distances) which move under the influence of their mutual gravitational attraction alone. The state of such a system at a given time is defined as determined by the positions and momenta of its component bodies at that time and does not include other aspects that might undergo change, such as the color or the chemical constitution of the moving bodies. The theory provides a set of laws—essentially, the Newtonian laws of gravitation and of motion—which, given the positions and momenta of the elements of such a system at any one time, mathematically determine their positions and momenta at any other time. In particular, those laws make it possible to offer a D-N explanation of the system's being in a certain state at a given time, by specifying, in the sentences $C_1, C_2, \ldots, C_k$ of the schema (D-N), the state of the system at some earlier time. The theory here referred to has been applied, for example, in accounting for the motions of planets and comets, and for solar and lunar eclipses.

In the explanatory or predictive use of a deterministic theory, then, the notion of a cause as a more or less narrowly circumscribed antecedent event has been replaced by that of some antecedent state of the total system, which provides the "initial conditions" for the computation, by means of the theory, of the later state that is to be explained. If the system is not isolated, i.e., if relevant outside influences act upon the system during the period of time from the initial state invoked to the state to be explained, then the particular circumstances that must be stated in the explanans include also those outside influences; and it is these "boundary conditions" in conjunction with the "initial" conditions which replace the everyday notion of cause, and which are specified by the statements $C_1, C_2, \ldots, C_k$ in the schematic representation (D-N) of deductive-nomological explanation.[20]

Causal explanation in its various degrees of explicitness and precision is not, however, the only mode of explanation on which the D-N model has a bearing. For example, the explanation of a general law by deductive subsumption under theoretical principles is clearly not an explanation by causes. But even when used to account for individual events, D-N explanations are not always causal. For example, the fact that a given simple pendulum takes two seconds to complete one full swing might be explained by pointing out that its length is 100 centimeters, and that the period t (in seconds), of any simple pendulum is connected with its length l (in centimeters) by the law that $t=2\pi\sqrt{l/g}$, where g is the acceleration of free fall. This law expresses a mathematical relationship between the length and the period (which is a quantitative dispositional characteristic) of the pendulum at one and the same time; laws of this kind, of which the laws of Boyle and of Charles, as well as Ohm's law are other examples, are sometimes called *laws of coexistence*, in contradistinction to *laws of succession*, which concern temporal changes in a system. These latter include, for example, Galileo's law and the laws for the changes of state in systems covered by a deterministic theory. Causal explanation by reference to antecedent events clearly presupposes laws of succession; in the case of the pendulum, where only a law of coexistence is invoked, one surely would not say that the pendulum's having a period of two seconds was *caused* by the fact that it had a length of 100 centimeters.

One further point deserves notice here. The law for the simple pendulum makes it possible not only to infer the period of a pendulum from its length, but also conversely to infer its length from its period; in either case, the inference is of the form (D-N). Yet a sentence stating the length of a given pendulum, in conjunction with the law, will be much more readily regarded as explaining the pendulum's period than a sentence stating the period, in conjunction with the law, would be considered as explaining the pendulum's length. This distinction appears to reflect the idea that we might change the length of the

20. For more detailed accounts of the notions of causality and of deterministic theory and deterministic system, see, for example, Feigl (1953); Frank (1957), chapters 11 and 12; Margenau (1950), chapter 19; Nagel (1961), pp. 73-78 and chapters 7 and 10.

pendulum at will and thus control its period as a "dependent variable," whereas the reverse procedure does not seem possible.[21] This conception is questionable, however; for we can also change the period of a given pendulum at will, namely, by changing its length. It cannot validly be argued that in the first case we have a change of length independently of a change of the period, for if the location of the pendulum remains fixed, then its length cannot be changed without also changing the period. In cases such as this, the common-sense conception of explanation appears to provide no clear grounds on which to decide whether a given argument that deductively subsumes an occurrence under laws is to qualify as an explanation.

In the instance just considered, a particular fact was explained, not by causal antecedents but by reference to another contemporaneous fact. It might even be argued that sometimes a particular event can be satisfactorily explained by reference to subsequent occurrences. Consider, for example, a beam of light that travels from a point *A* in one optical medium to a point *B* in another, which borders upon the first along a plane. Then, according to Fermat's principle of least time, the beam will follow a path that makes the traveling time from *A* to *B* a minimum as compared with alternative paths available. Which path this is will depend on the refractive indices of the two media; we will assume that these are given. Suppose now that the path from *A* to *B* determined by Fermat's principle passes through an intermediate point *C*. Then this fact may be said to be D-N explainable by means of Fermat's law in conjunction with the relevant data concerning the optical media and the information that the light traveled from *A* to *B*. But its "arrival at *B*," which thus serves as one of the explanatory factors, occurs only after the event to be explained, namely, the beam's passing through *C*.

Any uneasiness at explaining an event by reference to factors that include later occurrences might spring from the idea that explanations of the more familiar sort, such as our earlier examples, seem to exhibit the explanandum event as having been brought about by earlier occurrences; whereas no event can be said to have been brought about by factors some of which were not even realized at the time of its occurrence. Perhaps this idea also seems to cast doubt upon purported explanations by reference to simultaneous circumstances. But, while such considerations may well make our earlier examples of explanation, and all causal explanations, seem more natural or plausible, it is not clear what precise construal could be given to the notion of factors "bringing about" a given event, and what reason there would be for denying

21. In this connection, *cf.* the discussion of causal statements as recipes for producing a given effect, in Gasking (1955).

the status of explanation to all accounts invoking occurrences that temporally succeed the event to be explained.[22]

2.3 THE ROLE OF LAWS IN EXPLANATION. The D-N model, as we have seen, assigns to laws or theoretical principles the role of indispensable premises in explanatory arguments. I will now consider some alternative conceptions of the role of laws in explanation.

2.3.1. *The Conception of Laws as Inference Rules.* One recently influential view construes laws and theoretical principles as inference rules in accordance with which particular statements of empirical fact may be inferred from other such statements.

Thus Schlick once held the view, for which he gave credit to Wittgenstein, that "basically a natural law does not have the logical character of a 'proposition' but represents 'a direction for the formulation of propositions'."[23] Schlick espoused this idea largely because he held at the time that a genuine statement must be capable of strict verification by particular experiential findings—a requirement evidently not met by general laws, which pertain to indefinitely many particular cases. But the requirement of strict verifiability for sentences that are to qualify as empirically significant has long since been abandoned as too restrictive,[24] and it surely constitutes no good reason for construing laws as rules rather than as statements.

In a somewhat different vein, Ryle has characterized law statements as statements which are true or false, but which characteristically function as inference licenses authorizing inferential moves from the assertion of some factual statements to the assertion of others.[25] This conception has influenced the views of several others writers on the role of laws in scientific and historical explanation. Dray, for example, has offered some interesting considerations in support of it with special reference to historical explanation. He points out that since an explanation of a concrete historical event will usually have to take into account a large set of relevant factors, the corresponding covering law may well be so highly qualified as to possess only one single instance, namely, the

22. For further observations on this issue, *cf.* Scheffler (1957).

23. Schlick (1931), p. 190 of English translation. See also the discussion of this idea by Toulmin, who accepts it with certain qualifications (1953, pp. 90-105), and who develops, in a somewhat similar spirit, an extensive analogy between physical theories and maps (1953, chapter 4). For illuminating comments on Toulmin's views, and on the problem in general, see Nagel's review of Toulmin's book in *Mind* 63, pp. 403-12 (1954), reprinted in Nagel (1956), pp. 303-15.

24. For details, see the essay "Empiricist Criteria of Cognitive Significance: Problems and Changes," in this volume.

25. *Cf.* Ryle (1949), pp. 121-123 and Ryle (1950).

occurrence it explains. But under these circumstances, Dray questions the propriety of applying the term 'law', whose ordinary use "has 'other cases' built right into it."[26] He holds, therefore, that though, when offering the explanation 'E because $C_1, C_2, \ldots, C_n$', the historian "commits himself to the truth of the covering general statement, 'If $C_1 \ldots C_n$ then E', . . . the statement thus elicited . . . is surely nothing more than a formulation of the *principle of the historian's inference* when he says that from the set of factors specified, a result of this kind could reasonably be predicted. The historian's inference may be said to be *in accordance with* this principle. But it is quite another matter to say that his explanation entails a corresponding *empirical law*."[27] Dray conceives of such principles of inference as being "general hypotheticals" of the form 'if p then q'; and he holds that "to claim simply that a 'general hypothetical' lurks implicitly in the historian's explanation is to claim *considerably less* than covering law theorists generally do"; for if the general hypothetical is construed as an inference license in Ryle's sense, then "to say that the historian's explanation commits him to the covering 'law' is merely to say that it commits him . . . . to reasoning in a similar way in any further cases *which may turn up*, since he claims universal validity for the corresponding argument, 'p so q'."[28]

But surely, to claim universal validity for this argument scheme is to assert by implictaion, the general statement 'Whenever p then q', and vice versa: there is no difference in the strength of the claims, but only in the mode of expressing them. And if the general statement has only one instance, then so does the corresponding rule, and one might with equal justice question the propriety of qualifying the latter as a principle of inference, on the ground that the idea of such a principle or rule, no less than the idea of a law, carries with it a suggestion of generality.

In his remarks on the number of instances of a law, Dray seems to view a historical explanation as using only one general hypothetical, namely, in effect, a "minimum covering law" of the kind mentioned earlier. As a rule, however, an explanation will rely on a more or less comprehensive set of laws, each of which has many instances, and of which the narrower covering law is simply a highly specific consequence. But suppose that a given explanation does rely on just one highly specific generalization that has only one instance. Can that generalization be qualified as a law? Our discussion in section 2.1 bears on this question, and it will suffice to add here only a few brief remarks. Suppose that an attempt were made to explain Hitler's decision to invade Russia by means of the generalizatoin 'Anyone exactly like Hitler in all respects, and facing

26. Dray (1957), p. 40.

27. Dray (1957), p. 39. Italics the author's.

28. Dray (1957), p. 41. Italics supplied.

exactly the same circumstances, decides to invade Russia'. This clearly affords no explanation because the general statement invoked is equivalent to the sentence 'Hitler decided to invade Russia', which is not a general sentence at all, and which simply restates the explanandum; for being exactly like Hitler in all respects is the same thing as being identical with Hitler. Thus, the proposed generalization is nonlawlike because it is not essentially generalized.

But a general statement—such as one of the highly specific covering laws envisaged by Dray—may well have only one instance without being logically equivalent to a singular sentence. This feature, as we noted earlier, would not deprive the generalization of lawlike status and potential explanatory power.

The arguments here briefly considered, then, do not lend much support to the conception of laws and theoretical principles as rules or principles of inference. On the other hand, there are some considerations which clearly militate against this construal.

First, in the writings of scientists, laws and theoretical principles are treated as statements. For example, general statements are used in conjunction with singular statements about particular facts to serve as *premises* from which other statements about particular facts are inferred; similarly, statements of general form, such as laws of narrower scope, often appear as *conclusions* derived from more comprehensive laws. Again, general laws or theoretical principles are accepted or rejected on the basis of empirical tests in much the same way as statements of particular facts, such as those concerning the constitution of the earth's interior, for example.

Indeed—and this brings us to a second difficulty—the distinction here presupposed between singular sentences on the one hand and general sentences on the other has no precise meaning in reference to statements formulated in a natural language. For example, the statement that the earth is a sphere may be regarded as a singular sentence of the form '*Se*', which assigns to a particular object, the earth, a certain property, sphericity. But it may also be construed as a general statement, e.g., as asserting that there is a point in the interior of the earth from which all the points on its surface have the same distance. Similarly, the statement that a given crystal of salt is soluble in water may be construed as a singular statement ascribing solubility to a particular object, or, alternatively, as a statement of general character, asserting or implying that the given crystal will dissolve at any time upon being put into water.

A precise distinction of the kind here in question can be drawn if (i) the statements to be classified are expressed in a suitably formalized language that provides for quantificational notation, and (ii) every extra-logical term of the language is characterized either as primitive or as defined, each defined term possessing a unique definition in terms of primitives. A sentence of such a language may then be said to be essentially singular if it is logically equivalent to a sentence containing no defined terms and no quantifiers; all other sentences will be essentially general. The sentence 'The earth is spherical' will then be essentially singular if, for example, both 'the earth' and 'spherical' count as primitive terms of the language in which our statements are formulated; it will be essentially general if, for example, 'spherical' is defined by an expression containing one or more noneliminable quantifiers.

But even if we assume that a precise dividing line between singular and general statements has been drawn in this or a similar manner, the proposal to construe general statements as inference rules connecting singular statements still faces another, more serious difficulty: the formulation of law statements as inference rules proves difficult, if not impossible, and the resulting system of rules is awkward, to say the least. To be sure, a statement of the simple form 'All F are G', or '$(x)\ (Fx \supset Gx)$', where 'F' and 'G' are primitive predicates in the sense just explained, might be replaced by a rule licensing inferential transition from any sentence of the form 'Fi' (which is singular, i.e., quantifier-free) to the corresponding sentence of the form 'Gi'. But scientific explanations are often based on laws of a more complex structure; and for these, recasting in the form of inference rules connecting singular statements becomes problematic. Take the law, for example, that every metal has a specific melting point (at atmospheric pressure); i.e., that for every metal there exists a temperature T such that at any lower temperature and at no higher temperature the metal is solid at atmospheric pressure. The corresponding inference rule could not be construed as authorizing the transition from any sentence of the form 'i is a metal' to the sentence 'there is a temperature T such that at any lower temperature, but at no higher one, i is solid at atmospheric pressure'; for the conclusion thus obtained is not a sentence of singular form, but a statement involving both existential and universal quantifiers. Indeed, the subclauses 'at any temperature below T, i is solid' and 'at any temperature above T, i is nonsolid' have themselves the universal form of a law, and the general conception here under discussion would therefore seem to require that they in turn be construed as inference rules rather than as statements. But in the given context, this is not possible since they are qualified by the existential-quantifier phrase 'there is a temperature T such that. . . .'. In sum, the given law cannot be construed as tantamount to a rule establishing certain inferential connections among singular sentences. This is not to say that the law permits no such inferences: indeed, with its help (i.e., using it as an additional premise), we can infer from the statement 'this key is metal and is not liquid at 80°C and atmospheric pressure' further descriptive statements to the effect that the key won't be liquid at 74°C, 30°C, and other specific temperatures below 80°C, at atmospheric pressure. But these and similar inferential connections among singular statements which are mediated by the given law clearly do not exhaust its content; for, as we noted, the law also establishes connections, for example, between singular sentences ('*i* is a metal') and quantified ones ('there is a temperature *T* such that. . . .').

It may even happen that of two or more laws of complex form, none taken by itself establishes any inferential connections among singular sentences, whereas jointly they do. For example, two sentences of the form '$(x)\ [Fx \supset (\exists y)Rxy]$' and '$(x)[(\exists y)(Rxy) \supset Gx]$' jointly permit the inference from 'Fi' to 'Gi'; but, individually, neither of them establishes any connection among singular sentences. Thus, the totality of inferential transitions among singular sentences that are made possible by a set of laws or theoretical principles may far exceed the (logical- or class-) sum of the inferential connections established, among the same singular sentences, by the laws or theoretical principles individually. Hence, if one were to insist on construing scientific laws and theoretical principles as extralogical inference rules, licensing certain transitions among singular sentences, then one would have to do so, not for each of the laws and theoretical principles individually, but at once for the entire set of laws and principles assumed in a given context. No doubt the simplest way of doing this would be to formulate just one extralogical rule, authorizing all and only those transitions among singular statements which can be effected by using only purely logical rules of inference and by treating the laws and theoretical principles "as if" they were statements capable of functioning as additional premises in deductive arguments. But to adopt this rule would be simply to pay lip service to the construal of laws as rules rather than as statements.[29]

In sum, then, there is serious doubt, on purely logical grounds, whether *all* laws and theoretical principles can be adequately construed as inference rules. And even in the cases where this is possible, the preceding considerations

29. It is of interest to note here that Carnap, in his theory of logical syntax, explicitly provides for the possibility of constructing languages with extralogical rules of inference; see Carnap (1937), section 51. He calls the latter physical rules or P-rules. But he does not claim that all general laws or theoretical principles can be construed as such rules; and he emphasizes that the extent to which P-rules are to be countenanced in constructing a language will be a matter of convenience. For example, if we use P-rules, then the discovery of empirical phenomena that "conflict" with our previously accepted theories may oblige us to alter the rules of inference, and thus the entire formal structure, of our scientific language; whereas in the absence of *P*-rules, only a modification of some previously accepted theoretical statements is called for. W. Sellars (1953), (1958), also has advocated the admission of material rules of inference in connection with his analysis of subjunctive conditionals.

For a lucid survey and critical appraisal of various reasons that have been adduced in support of construing general laws as inference rules see Alexander (1958).

suggest that it would be simpler and more helpful, for a clarification of the issues with which we are here concerned, to construe general laws and theoretical principles as statements: hence this course will be followed from here on.

2.3.2 *The Conception of Laws as Role-Justifying Grounds for Explanations.* Another conception that would normally preclude the mention of laws in an explanation has been set forth by Scriven,[30] who argues that in so far as laws are relevant to an explanation, they will usually function as "role-justifying grounds" for it. This conception doubtless reflects the view that, as Ryle has put it, "Explanations are not arguments but statements. They are true or false."[31] Explanations might then take the form 'q because p', where the 'p'-clause mentions particular facts but no laws; and the kind of explanation represented as an argument in our schema (D-N) would be expressed by a statement of the form 'E because $C_1, C_2, \ldots, C_k$.' The citation of laws is appropriate, according to Scriven, not in response to the question 'Why q?', which 'q because p' serves to answer, but rather in response to the quite different question as to the grounds on which the facts mentioned in the 'p'-clause may be claimed to explain the facts referred to in the 'q'-clause. To include the relevant laws in the statement of the explanation itself would be, according to Scriven, to confound the statement of an explanation with a statement of its grounds.

Now it is quite true that in ordinary discourse and also in scientific contexts, a question of the form 'Why did such-and-such an event happen?' is often answered by a because-statement that cites only certain particular facts—even in cases where the relevant laws could be stated. The explanation statement 'The ice cube melted because it was floating in water at room temperature' is an example. But as this sentence equally illustrates, an explanation as ordinarily formulated will often mention only some of a larger set of particular facts which jointly could explain the occurrence in question. It will forego mention of other factors, which are taken for granted, such as that the water as well as the surrounding air remained approximately at room temperature for an adequate time. Hence, in order to justify attributing an explanatory role to the facts actually specified, one would have to cite here not only certain laws, but also the relevant particulars that had not been explicitly mentioned among the explanatory facts. Thus it is not clear why only laws should be singled out for the function of role-justification.[32] And if statements of particular fact were equally allowed to serve as role-justifying grounds in explanations, then the distinction between explanatory facts and role-justifying grounds would become obscure and arbitrary.

30. Scriven (1959), especially section 3.1.
31. Ryle (1950), p. 330.
32. The same point has been made by Alexander (1958, section I).

Scriven goes beyond relegating explanatory laws to the place of role-justifying grounds: He holds that we can sometimes be quite certain of a given explanation without being able to justify it by reference to any laws; in his own words, "certain evidence is adequate to guarantee certain explanations without the benefit of deduction from laws."[33] One of his examples is this:

> As you reach for the dictionary, your knee catches the edge of the table and thus turns over the ink-bottle, the contents of which proceed to run over the table's edge and ruin the carpet. If you are subsequently asked to explain how the carpet was damaged you have a complete explanation. You did it, by knocking over the ink. The certainty of this explanation is primeval. It has absolutely nothing to do with your knowledge of the relevant laws of physics; a cave-man could supply the same account and be quite as certain of it. . . . If you were asked to produce the role-justifying grounds for your explanation, what could you do? *You could not produce any true universal hypothesis* in which the antecedent was identifiably present (i.e., which avoids such terms as "knock hard enough"), and the consequent is the effect to be explained.[34]

At best, Scriven continues, one could offer a vague generalization to the effect that if you knock a table hard enough, it will cause an ink-bottle not too securely placed on it to spill over provided that there is enough ink in it. But this needs tightening in many ways, and, Scriven claims, it cannot be turned into a true universal hypothesis which, for the example in question, would "save the deductive model." In particular, physics cannot be expected to yield such a hypothesis, for "the explanation has become not one whit more certain since the laws of elasticity and inertia were discovered."[35]

Undeniably, in our everyday pursuits and also in scientific discussions, we often offer or accept explanatory accounts of the sort illustrated by Scriven's example. But an analytic study of explanation cannot content itself with simply registering this fact: it must treat it as material for analysis; it must seek to clarify what is *claimed* by an explanatory statement of this sort, and how the claim might be *supported*. And, at least to the first question, Scriven offers no explicit answer. He does not tell us just what, on his construal, is asserted by the given law-free explanation; and it remains unclear, therefore, precisely what claim he regards as having primeval certainty, for cave-man and modern physicist alike. Presumably the explanation he has in mind would be expressed by a statement roughly to the effect that the carpet was stained with ink because the table was knocked. But, surely, this statement claims by implication that the antecedent circumstances invoked were of a kind which generally yields effects of the sort to be explained. Indeed, it is just this implicit claim

33. Scriven (1959), p. 456.
34. *Loc. cit.*, italics the author's.
35. *Loc. cit.*

of covering uniform connections which distinguishes the causal attribution here made from a mere sequential narrative to the effect that first the table was knocked, then the bottle tipped over, and finally the ink dripped on the rug. Now, in a case such as the spilling of the ink, we feel familiar, at least in a general manner, with the relevant uniform connections even though we may not be able to state them precisely, and thus we are willing to take them for granted without explicit mention. On the other hand, there are various conceivable, particular antecedents any one of which might, by virtue of roughly the same general uniformities, account for the tipping over of the ink bottle: I might have knocked the table, the cat might have pushed the ink bottle, the curtain might have brushed against the bottle in a breeze, and so forth. Thus, the question of how the ink spot got on the rug will usually be aimed at eliciting information about the particular antecedents that led to the damage; and it might seem, therefore, that an explanation need have nothing to do with uniformities or laws. But this appearance surely does not refute the view that any particular explanatory claim made in terms of antecedent circumstances still presupposes suitable covering laws.

This brings us to a crucial question posed by Scriven's argument. Is it possible to specify, in the given case, a set of laws which would actually provide role-justification, by enabling us to deduce the explanandum, given the information about the antecedent explanatory events? The question cannot be answered unequivocally because it is too vague. Assuming that the explanatory statement takes the form '*q* because *p*', we have not been told precisely what takes the places of '*p*' and of '*q*' in the case of the overturned ink bottle. If, for example, the '*p*'-statement were taken to include the information that a full, uncorked, ink bottle was in fact knocked over, and if the '*q*'-statement reported merely that the ink leaked out, then some elementary laws in the mechanics of fluids might well provide adequate nomological support for the explanatory statement. If, by contrast, the '*q*'-statement is taken to specify, not only that the ink spilled out, but also that it produced a stain of specified size and shape on the rug, then, to be sure, no laws are known that would permit the inference from the '*p*'-statement (in any plausible construal) to this '*q*'-statement. But, just for this reason, an account of the sort suggested by Scriven's example would not be regarded as explaining the size or the shape of the ink stain at all.

No doubt, the explanatory claim envisaged by Scriven lies between these extremes and is roughly to the effect that the rug was stained because the table, with an open bottle of ink standing on it, was caught and lifted by my knee. This claim might be paraphrased by saying that there are laws connecting the presence of an ink stain on the rug with certain antecedent circumstances, which include an open bottle of ink standing on the table, and the fact that the table's edge was lifted. And there seems to be no reason to doubt the possibility of adducing or establishing a gradually expanding set of laws which would afford an increasingly accurate and detailed explanation of the phenomenon at hand.

We might say, in agreement with Scriven, that these laws would lend support or justification to the given because-statement. But we should note also that an expansion of the set of supporting laws will normally call for a corresponding expansion of the set of antecedent circumstances which have to be taken into account, and thus, strictly, for a modification of the explanatory because-statement itself.

Furthermore, the task of establishing the statements, whether of laws or of particular facts, which may thus be invoked in support of a because-statement comes clearly within the domain of scientific inquiry; hence it cannot reasonably be argued that progress in physical or chemical reaserch has no significance for the explanation at hand. Thus Scriven's cave man, or perhaps a child, might well assume that when any opaque liquid is poured on any kind of textile it will soak in and produce a stain; which would lead him to expect a stain when mercury is dropped on a rug or when ink is poured on a specially treated nonstaining textile. And if his explanation or understanding of the ink stain on the rug presupposes that assumption then it would plainly be far from primevally certain: it would be false.

In sum then, the claim that the cave man could explain the staining of the rug with the same "certainty" as a modern scientist loses its initial striking plausibility when we ask ourselves precisely what the explanation would assert and what it would imply, and when we make sure it is not simply taken to be a narration of selected stages in the process concerned. An explanation may well be put into the form of a sequential narrative, but it will explain only if it at least tacitly presupposes certain nomic connections between the different stages cited. Such "genetic" explanations will be examined more closely later in this essay.

In the preceding discussion we have construed an explanatory statement of the form 'q because p' as an assertion to this effect: p is (or was) the case, and there are laws (not explicitly specified) such that the statement that q is (or was) the case follows logically from those laws taken in conjunction with the statement of p and perhaps other statements, which specify antecedents not included in p but tacitly presupposed in the explanation. In his discussion of the explanatory role of laws, Scriven considers the closely related idea that when we are able to specify the cause of a particular event such as the staining of the rug, "we are in a position to judge, not that certain specifiable laws apply, but that *some* laws must apply." And he objects that "it is *very* odd to say this rather than that we can sometimes be quite sure of causal statements even when we do not know any relevant laws. This capacity for identifying causes is learnt, is better developed in some people than in others, can be tested, and is the basis for what we call *judgments*."[36]

But this surely is no telling objection. For first of all, if the thesis is to have a clear meaning we need to know exactly what is meant by 'identifying the cause of a particular event', and how, accordingly, the capacity for identifying causes may be tested: and Scriven does not provide this information.

Secondly, the conception that a statement of the form '*q* because *p*' asserts, by implication, the existence of certain covering laws is by no means incompatible with the view that people may have a capacity for causal judgment even when they are unable to specify suitable covering laws or to explicate the notion of cause they are using. Consider a parallel: An experienced carpenter or gardener may have a capacity for judging very accurately the size of the area enclosed by a given circular line without being able to give an analytic definition of the area of a circle in terms of the convergent series formed by the areas of certain inscribed or circumscribed polygons. But this surely would not justify the claim that therefore, at least in the specific cases accessible to the judgments of skilled craftsmen, the mathematical analysis of the concept of the area of a circle is irrelevant or does not apply. Similarly a physician, a garage mechanic, or an electrician may have a remarkable capacity for judging what causes trouble in a particular case without always being able to adduce general laws supporting the diagnosis, and indeed without even believing that the latter presupposes the existence of such laws. But this acknowledgment does not warrant the conclusion that it is impossible or inappropriate to construe the causal statements in question as making reference to, or at least implying the existence of, corresponding laws.

Even the way in which causal statements based on such practical "judgment" are tested and substantiated indicates that they make, at least implicitly, a claim of general character. Thus, the assertion that a certain therapeutic measure caused improvement in a given case would require corroboration by similar results in similar cases, so as to rule out the possibility of a mere coincidence as contradistinguished from a causal connection.

But, since explanatory accounts are often formulated as 'because'-statements, should we not at least introduce a further model, which construes explanations as statements of the form '*q* because *p*' rather than as arguments? To characterize a certain type of explanation simply as having that form would surely be

36. *Loc. cit.*, italics the author's.

insufficient: the chief task of the contemplated model would be to clarify the meaning of the word 'because' in explanatory contexts, and this requires further analysis. To claim that we can sometimes proffer explanations of the form '*q* because *p*' with complete certainty, or that they can be guaranteed by suitable kinds of evidence without the benefit of laws, is to sidestep this issue; indeed, the claim cannot even be assessed independently of an analysis of the explanatory use of the word 'because'. The paraphrasing of because-statements suggested above is rather vague and no doubt capable of improvement, but at least it seems to me correct in exhibiting the assumption of lawlike connections implicit in such explanatory formulations.

2.4 EXPLANATION AS POTENTIALLY PREDICTIVE. Because of its essential reliance on laws and theoretical principles, D-N explanation may be expected to show a close affinity to scientific prediction; for laws and theoretical principles, making general claims, range also over cases not as yet examined and have definite implications for them.

The affinity in question is vividly illustrated in the fourth part of the *Dialogues Concerning Two New Sciences.* Here, Galileo develops his laws for the motion of projectiles and deduces from them the corollary that if projectiles are fired from the same point with equal initial velocity, but different elevations, the maximum range will be attained when the elevation is 45°. Then, Galileo has Sagredo remark: "From accounts given by gunners, I was already aware of the fact that in the use of cannon and mortars, the maximum range. . . is obtained when the elevation is 45°. . .; but to understand why this happens far outweighs the mere information obtained by the testimony of others or even by repeated experiment."[37] The reasoning that affords such understanding can readily be put into the form (D-N); it amounts to a deduction, by logical and mathematical means, of the corollary from a set of premises that contains (i) the fundamental laws of Galileo's theory for the motion of projectiles and (ii) particular statements specifying that all the missiles considered are fired from the same place with the same initial velocity. Clearly, then, the phenomenon previously noted by the gunners is here *explained*, and thus *understood*, by showing that its occurrence was to be expected under the specified circumstances in view of certain general laws set forth in Galileo's theory. And Galileo himself points with obvious pride to the *predictions* that may in like fashion be obtained by deduction from his laws; the latter imply "what has perhaps never been observed in experience, namely, that of other shots those which exceed or fall short of 45° by equal amounts have equal ranges." Thus, the explanation

37. Galilei (1946), p. 265.

afforded by Galileo's theory "prepares the mind to understand and ascertain other facts without need of recourse to experiment,"[38] namely, by deductive subsumption under the laws on which the explanation is based.

Checking the predictions thus derived from the general laws or theoretical principles invoked in an explanation is an important way of testing those "covering" generalizations, and a favorable outcome may lend strong support to them. Consider, for example, the explanation offered by Torricelli for a fact that had intrigued his teacher Galileo; namely, that a lift pump drawing water from a well will not raise the water more than about 34 feet above the surface of the well.[39] To account for this, Torricelli advanced the idea that the air above the water has weight and thus exerts pressure on the water in the well, forcing it up the pump barrel when the piston is raised, for there is no air inside to balance the outside pressure. On this assumption the water can rise only to the point where its pressure on the surface of the well equals the pressure of the outside air on that surface, and the latter will therefore equal that of a water column about 34 feet high.

The explanatory force of this account hinges on the conception that the earth is surrounded by a "sea of air" that conforms to the basic laws governing the equilibrium of liquids in communicating vessels. And because Torricelli's explanation presupposed such general laws it yielded predictions concerning as yet unexamined phenomena. One of these was that if the water were replaced by mercury, whose specific gravity is about 14 times that of water, the air should counterbalance a column about 34/14 feet, or somewhat less than $2\frac{1}{2}$ feet, in length. This prediction was confirmed by Torricelli in the classic experiment that bears his name. In addition, the proposed explanation implies that at increasing altitudes above sea level, the length of the mercury column supported by air pressure should decrease because the weight of the counterbalancing air decreases. A careful test of this prediction was performed at the suggestion of Pascal only a few years after Torricelli had offered his explanation: Pascal's brother-in-law carried a mercury barometer (i.e., essentially a mercury column counterbalanced by the air pressure) to the top of the Puy-de-Dôme, measuring the length of the column at various elevations during the ascent and again during the descent; the readings were in splendid accord with the prediction.[40]

The inferences by which such predictions are obtained are again of deductive-

38. *Loc. cit.*

39. The following account is based on the presentation of this case in Conant (1951), chapter 4.

40. Pascal's own account and appraisal of the "great experiment" is reprinted in English translation in Moulton and Schifferes (1945), pp. 145-53.

nomological form: The premises comprise the explanatory laws in question (in our last example, especially Torricelli's hypothesis) and certain statements of particular fact (e.g., that a barometer of such and such construction will be carried to the top of a mountain). Let us refer to predictive arguments of the form (D-N) as D-N predictions. In empirical science many predictive arguments are of this kind. Among the most striking examples are forecasts, based on the principles of celestial mechanics and of optics, concerning the relative positions of the Sun, the Moon, and the planets at a given time, and concerning solar and lunar eclipses.

It may be well to stress here that while the principles of classical mechanics or other deterministic laws or theories afford the basis for very impressive D-N explanations and predictions, the additional premises required for this purpose must provide not only a specification of the state of the system at some time t_0 earlier than the time t_1 for which the state of the system is to be inferred, but also a statement of the boundary conditions prevailing between t_0 and t_1; these specify the external influences acting upon the system during the time interval in question. For certain purposes in astronomy the disturbing influence of celestial objects other than those explicitly considered may be neglected as insignificant, and the system under consideration may be treated as "isolated"; but this should not lead us to overlook the fact that even those exemplars of deductive-nomological prediction do not enable us to forecast future events strictly on the basis of information about the present: the predictive argument also requires certain premises concerning the future—e.g., absence of disturbing influences, such as a collision of Mars with an unexpected comet; and the temporal scope of these boundary conditions must extend up to the very time of occurrence of the predicted event. The assertion therefore that laws and theories of deterministic form enable us to predict certain aspects of the future from information about the present has to be taken with a grain of salt. Analogous remarks apply to deductive-nomological explanation.

Since in a fully stated D-N explanation of a particular event the explanans logically implies the explanandum, we say may that the explanatory argument might have been used for a deductive prediction of the explanandum-event *if* the laws and the particular facts adduced in its explanans had been known and taken into account at a suitable earlier time. In this sense, a D-N explanation is a potential D-N prediction.

This point was made already in an earlier article by Oppenheim and myself,[41] where we added that scientific explanation (of the deductive-nomological kind) differs from scientific prediction not in logical structure, but in certain

41. Hempel and Oppenheim (1948), section 3.

pragmatic respects. In one case, the event described in the conclusion is known to have occurred, and suitable statements of general law and particular fact are sought to account for it; in the other, the latter statements are given and the statement about the event in question is derived from them before the time of its presumptive occurrence. This conception, which has sometimes been referred to as the *thesis of the structural identity* (or of the symmetry) *of explanation and prediction*, has recently been questioned by several writers. A consideration of some of their arguments may help to shed further light on the issuse involved.

To begin with, some writers[42] have noted that what is usually called a prediction is not an argument but a sentence. More precisely, as Scheffler has pointed out, it is a sentence-token, i.e., a concrete utterance or inscription of a sentence purporting to describe some event that is to occur after the production of the token.[43] This is certainly so. But in empirical science predictive sentences are normally established on the basis of available information by means of arguments that may be deductive or inductive in character; and the thesis under discussion should be understood, of course, to refer to explanatory and predictive *arguments*.

Thus construed, *the thesis of structural identity amounts to* the conjunction of *two sub-theses*, namely (i) that *every adequate explanation is potentially a prediction* in the sense indicated above; (ii) that conversely *every adequate prediction is potentially an explanation*. I will now examine a number of objections that have been raised against the thesis, dealing first with those which, in effect, concern the first sub-thesis, and then with those concerning the second sub-thesis. I will argue that the first sub-thesis is sound, whereas the second one is indeed open to question. Though the following considerations are concerned principally with D-N explanation, some of them are applicable to other types of explanation as well. The adequacy of the structural identity thesis for the case of statistical explanation will be examined in detail in section 3.5.

The first sub-thesis, as has already been noted, is an almost trivial truth in the case of D-N explanation, since here the explanans logically implies the explanandum. But it is supported also by a more general principle, which applies to other types of explanation as well, and which expresses, I would submit, a general *condition of adequacy for any rationally acceptable explanation of a particular event*. That condition is the following: Any rationally acceptable answer to the question 'Why did event *X* occur?' must offer information

42. See Scheffler (1957), section 1 and (1963), Part I, sections 3 and 4; Scriven (1962), p. 177.

43. *Cf.* Scheffler (1957), section 1. For a more detailed study of explanation and prediction in the light of the type-token distinction, see Kim (1962).

which shows that *X* was to be expected—if not definitely, as in the case of D-N explanation, then at least with reasonable probability. Thus, the explanatory information must provide good grounds for believing that *X* did in fact occur; otherwise, that information would give us no adequate reason for saying: "That explains it—that does show why *X* occurred." And an explanatory account that satisfies this condition constitutes, of course, a potential prediction in the sense that it could have served to predict the occurrence of *X* (deductively or with more or less high probability) if the information contained in the explanans had been available at a suitable earlier time.

The condition of adequacy just stated can be extended, in an obvious manner, to explanations concerned, not with individual events, but with empirical uniformities expressed by putative laws. But such explanations cannot well be spoken of as potential *predictions* since law-statements purport to express timeless uniformities and thus make no reference to any particular time, whether past, present, or future.[44]

It will hardly be necessary to emphasize that it is not, of course, the *purpose* of an explanation to provide grounds in support of the explanandum-statement; for, as was noted in the first section of this essay, a request for an explanation normally *presupposes* that the explanandum-statement is true. The point of the preceding remarks is rather that an adequate explanation cannot help providing information which, if properly established, also provides grounds in support of the explanandum-statement. In the terminology of section 1, we may say that an adequate answer to an explanation-seeking why-question is always also a potential answer to the corresponding epistemic why-question.

The converse, however, does not hold; the condition of adequacy is necessary but not sufficient for an acceptable explanation. For example, certain empirical findings may give excellent grounds for the belief that the orientation of the earth's magnetic field shows diurnal and secular variations, without in the least explaining why. Similarly, a set of experimental data may strongly *support* the assumption that the electric resistence of metals increases with their temperature or that a certain chemical inhibits the growth of cancer cells, without providing any *explanation* for these presumptive empirical regularities. The predictive inferences here involved are inductive rather than deductive; but what bars them from the status of potential explanations is not their inductive character (in section 3, we will deal with inductive arguments that afford perfectly good scientific explanations), but the fact that they invoke no laws or theoretical principles, no explanatory statements that make a general claim. Reliance on general connecting principles, while not indispensable for

44. This point is made, for example, by Scriven (1962), pp. 179ff.

prediction, is required in any explanation: such principles alone can give to whatever particular circumstances may be adduced the status of explanatory factors for the event to be explained.

Some of the objections recently raised against the thesis of the structural identity of explanation and prediction concern in effect the first of its two sub-theses, which has now been presented in some detail: the claim that any adequate explanatory argument is also potentially predictive. I will consider three objections to the effect that there are certain perfectly satisfactory explanations that do not constitute potential predictions.

Scriven has argued that the occurrence of an event X is sometimes quite adequately explained by means of a "proposition of the form 'The only cause of X is A' . . . for example, 'The only cause of paresis is syphilis';" this proposition enables us to explain why a certain patient has paresis by pointing out that he previously suffered from syphilis. And this explanation holds good, according to Scriven, even though only quite a small percentage of syphilitic patients develop paresis, so that "we must, on the evidence [that a given person has syphilis], still predict that [paresis] will *not* occur."[45] But if it does occur, then the principle that the only cause of paresis is syphilis can "provide and guarantee our explanation" in terms of antecedent syphilitic infection.[46] Thus we have here a presumptive explanation which indeed is not adequate as a potential prediction. But precisely because paresis is such a rare sequel of syphilis, prior syphilitic infection surely cannot by itself provide an adequate explanation for it. A condition that is nomically necessary for the occurrence of an event does not, in general, explain it; or else we would be able to explain a man's winning the first prize in the Irish sweepstakes by pointing out that

45. Scriven (1959a), p. 480, italics the author's.

46. *Loc. cit.* Barker has argued analogously that "it can be correct to speak of explanation in many cases where specific prediction is not possible. Thus, for instance, if the patient shows all the symptoms of pneumonia, sickens and dies, I can then explain his death—I know what killed him—but I could not have definitely predicted in advance that he was going to die; for usually pneumonia fails to be fatal." (1961, p. 271). This argument seems to me open to questions similar to those just raised in reference to Scriven's illustration. First of all, it is not clear just what would be claimed by the assertion that pneumonia killed the patient. Surely the mere information that the patient had pneumonia does not suffice to explain his death, precisely because in most cases pneumonia is not fatal. And if the explanans is taken to state that the patient was suffering from very severe pneumonia (and perhaps that he was elderly or weak) then it may well provide a basis at least for a probabilistic explanation of the patient's death—but in this case it obviously also permits prediction of his death with the same probability. For some further observations on Barker's argument, see the comments by Feyerabend and by Rudner, and Barker's rejoinders, in Feigl and Maxwell (1961), pp. 278-85. A detailed critical discussion that sheds further light on Scriven's paresis example will be found in Grünbaum (1963) and (1963a), chapter 9; see also Scriven's rejoinder (1963).

he had previously bought a ticket, and that only a person who owns a ticket can win the first prize.

A second argument which, like Scriven's, has considerable initial plausibility has been advanced by Toulmin[47] by reference to "Darwin's theory, explaining the origin of species by variation and natural selection. No scientist has ever used this theory to foretell the coming-into-existence of creatures of a novel species, still less verified his forecast. Yet many competent scientists have accepted Darwin's theory as having great explanatory power." In examining this argument, let me distinguish what might be called the *story* of evolution from the *theory* of the underlying mechanisms of mutation and natural selection. The story of evolution, as a hypothesis about the gradual development of various types of organisms, and about the subsequent extinction of many of these, has the character of a hypothetical historical narrative *describing* the putative stages of the evolutionary process; it is the associated theory which provides what *explanatory insight* we have into this process. The story of evolution might tell us, for example, that at a certain stage in the process dinosaurs made their appearance and that, so much later, they died out. Such a narrative account does not, of course, explain why the various kinds of dinosaurs with their distinctive characteristics came into existence, nor does it explain why they became extinct. Indeed even the associated theory of mutation and natural selection does not answer the first of these questions, though it might be held to shed some light on the latter. Yet, even to account for the extinction of the dinosaurs, we need a vast array of additional hypotheses about their physical and biological environment and about the species with which they had to compete for survival. But if we have hypotheses of this kind that are specific enough to provide, in combination with the theory of natural selection, at least a probabilistic explanation for the extinction of the dinosaurs, then clearly the explanans adduced is also qualified as a basis for a potential probabilistic prediction. The undeniably great persuasiveness of Toulmin's argument would seem to derive from two sources, a widespread tendency to regard the basically descriptive story of evolution as explaining the various states of the process, and a similarly widespread tendency to overestimate the extent to which even the theory of mutation and natural selection can account for the details of the evolutionary sequence.

I now turn to a third objection to the claim that an adequate explanation is also a potential prediction. It is based on the observation that sometimes the only ground we have for asserting some essential statement in the explanans lies

47. Toulmin (1961), pp. 24–25. Scriven (1959a) and Barker (1961) have offered arguments in the same vein. For a critical discussion of Scriven's version, see Grünbaum (1963) and (1963a), chapter 9.

in the knowledge that the explanandum event did in fact occur. In such cases, the explanatory argument clearly could not have been used to predict that event. Consider one of Scriven's examples.[48] Suppose that a man has killed his wife whom he knew to have been unfaithful to him, and that his action is explained as the result of intense jealousy. The fact that the man was jealous might well have been ascertainable before the deed, but to explain the latter, we need to know that his jealousy was intense enough to drive him to murder; and this we can know only after the deed has actually been committed. Here then, the occurrence of the explanandum event provides the only grounds we have for asserting one important part of the explanans; the explanandum event could not therefore have been predicted by means of the explanatory argument. In another example,[49] Scriven considers an explanation to the effect that the collapse of a bridge was caused by metal fatigue. This account, he argues, might be supported by pointing out that the failure could have been caused only by an excessive load, by external damage, or by metal fatigue, and that the first two factors were not present in the case at hand, whereas there is evidence of metal fatigue. *Given the information that the bridge did in fact collapse*, this would establish not only that metal fatigue was at fault but that it was strong enough to cause the failure. While Scriven's notion of "the only possible cause" of a given event surely requires further elucidation, his example does afford another illustration of an explanatory account one of whose constituent hypotheses is supported only by the occurrence of the event to be explained—so that the latter could not have been predicted by means of the explanatory argument.

However, the point thus illustrated does not affect at all the conditional thesis that an adequate explanatory argument must be such that it could have served to predict the explanandum event *if* the information included in the explanans had been known and taken into account before the occurrence of that event. What Scriven's cases show is that sometimes we do not know independently of the occurrence of the explanandum event that all the conditions listed in the explanans are realized. However, this means only that in such cases our conditional thesis is counterfactual, i.e., that its if-clause is not satisfied, but not that the thesis itself is false. Moreover, Scriven's argument does not even show that in the kind of case he mentions it is logically or nomologically impossible (impossible by reason of the laws of logic or the laws of nature) for us to know the critical explanatory factor before, or independently of, the occurrence of the explanandum-event; the impossibility appears to be rather a practical and perhaps temporary one, reflecting present limitations of knowledge or technology.

48. Scriven (1959), pp. 468-69.
49. Scriven (1962), pp. 181-87.

But while it thus leaves our thesis unaffected, Scriven's observation is of methodological interest in its own right: it shows that sometimes an event is explained by means of hypotheses for some of which the fact of its occurrence affords the only available evidential support. This may happen, as we saw, when one of the explanatory hypotheses states that a certain relevant factor was strong enough to bring about the event in question; but the observation applies also to other cases. Thus the explanation, outlined in section 2.1, of the appearance and initial growth of the soap bubbles, includes in its explanans the assumption that a soap film had formed between the plate and the rims of the tumblers; and practically the only evidence available in support of this explanatory assumption is the fact that soap bubbles did emerge from under the tumblers. Or consider the explanation of the characteristic dark lines in the absorption spectrum of a particular star. The key assumption in the explanans is that the star's atmosphere contains certain elements, such as hydrogen, helium, and calcium, whose atoms absorb radiation of the wave lengths corresponding to the dark lines; the explanation relies, of course, on many other assumptions, including the optical theory that forms the basis for spectroscopy, and the assumption that the apparatus used is a properly constructed spectroscope. But while these latter explanans statements are capable of independent test and corroboration, it may well be that the only evidence available in support of the key explanatory hypothesis is the occurrence of the very lines whose appearance in the spectrum the argument serves to explain. Strictly speaking, the explanandum event here provides support for the key explanatory hypothesis only by virtue of the background theory, which connects the presence of certain elements in the atmosphere of a star with the appearance of corresponding absorption lines in its spectrum. Thus, the information that the explanandum event has occurred does not by itself support the explanatory hypothesis in question, but it constitutes, as we might say, an essential part of the only evidence available in support of that hypothesis.

Explanations of the kind here considered may be schematically characterized as arguments of the form (D-N) in which the information or assumption that E is true provides an indispensable part of the only available evidential support for one of the explanans statements, say, C_1. Let us call such explanations *self-evidencing*. It might be held that the actual occurrence of the explanandum event always provides some slight additional support even for an explanans whose constituent sentences have been accepted on the basis of independent evidence, and that in this sense every D-N explanation with true explanandum is in some measure self-evidencing; but we will apply this appellation to an explanatory account only if, at the time of its presentation, the occurrence of the explanandum event provides the only evidence, or an indispensable part of the only evidence, available in support of some of the explanans-statements.

An explanatory argument of the form (D-N) which is self-evidencing is not for that reason circular or pointless. To be sure, if the same argument were adduced in support of the assertion that the explanandum-event did occur (or, that E is true), then it would be open to the charge of epistemic circularity. If the argument is to achieve its objective then all the grounds it adduces in support of E—i.e., $C_1, C_2, \ldots, C_k; L_1, L_2, \ldots, L_r$—would have to be established independently of E; and this condition is violated here since the only ground we have for believing or asserting C_1 includes the assumption that E is true. But when the same argument is used for explanatory purposes it does not claim to establish that E is true; that is *presupposed* by the question 'Why did the event described by E occur?'. Nor need a self-evidencing explanation involve an explanatory circle. The information that the explanandum event has occurred is not included in the explanans (so that the occurrence of the event is not "explained by itself"); rather it serves, quite outside the explanatory context, as evidence supporting one of the explanans statements. Thus, an acceptable self-evidencing explanation benefits, as it were, by the wisdom of hindsight derived from the information that the explanandum event has occurred, but it does not misuse that information so as to produce a circular explanation.

An explanation that is self-evidencing may for that reason rest on a poorly supported explanans and may therefore have no strong claim to empirical soundness. But even this is not inevitable. In the case of the absorption spectrum of a star, for example, the previously accepted background information, including the relevant theories, may indicate that the dark lines observed occur *only* if the specified elements are present in the star's atmosphere; and then the explanandum, in conjunction with the background information, lends very strong support to the crucial explanatory hypothesis.

The notion of a self-evidencing explanation can, I think, shed some further light on the puzzle illustrated by the explanation of paresis in terms of antecedent syphilitic infection. Consider another illustration. Some cases of skin cancer are attributed to intensive ultraviolet irradiation. But this factor very often does not lead to cancer, so that the information that a person has been exposed to such radiation does not permit the prediction of cancer. Is that information alone nevertheless sufficient to explain the development of skin cancer when it does follow intensive irradiation? No doubt, an explanation will often be formulated so as to mention only the antecedent irradiation; but the underlying rationale surely must be more complex. Leaving aside the important quantitative aspects of the problem, the crucial point in that rationale can, I suggest, be schematically stated as follows: Some, though by no means all, individuals have the disposition to develop skin cancer upon exposure to strong ultraviolet irradiation; let us call these radiation-sensitive. Now, in the case of explanation, we know that the given individual was exposed to strong radiation (C_1) and did develop cancer of the skin in the affected area (E). But jointly, these two pieces of information lend support to the assumption that the individual is radiation-sensitive (C_2)—an hypothesis that is not supported in the case of prediction, where C_1 is available, but not E. And the two statements C_1 and C_2 (in combination with the general statement that sensitive individuals will develop skin cancer when exposed to intensive radiation) do provide an adequate explanans for E. If the explanation is thus construed as invoking C_2 in addition to C_1, it is seen to be self-evidencing, but also to possess an explanans which would provide an adequate basis for prediction if C_2 could be known in advance. That is impossible, of course, as long as the only available test for radiation-sensitivity consists in checking whether an individual does develop skin cancer upon intensive irradiation. But, clearly, it is conceivable that other, independent, tests of radiation-sensitivity might be found and then C_2 might well be established independently of, and even prior to, the occurrence of the event described by E.

In discussing the structural identity of explanation and prediction, I have so far considered only the first of the two sub-theses distinguished earlier, namely, the claim that every adequate explanation is also a potential prediction. I have argued that the objections raised against this claim fall short of their mark, and that the first sub-thesis is sound and can indeed serve as a necessary condition of adequacy for any explicitly stated, rationally acceptable explanation.

I turn now to the second sub-thesis, namely, that every adequate predictive argument also affords a potential explanation. This claim is open to question even in the case of certain predictive arguments that are of deductive-nomological character, as the following example illustrates. One of the early symptoms of measles is the appearance of small whitish spots, known as Koplik spots, on the mucous linings of the cheeks. The statement, L, that the appearance of Koplik spots is always followed by the later manifestations of the measles might therefore be taken to be a law, and it might then be used as a premise in D-N arguments with a second premise of the form 'Patient i has Koplik spots at time t', and with a conclusion stating that i subsequently shows the later manifestations of the measles. An argument of this type is adequate for predictive purposes, but its explanatory adequacy might be questioned. We would not want to say, for example, that i had developed high fever and other symptoms of the measles because he had previously had Koplik spots.

Yet this case—and others similar to it—does not constitute a decisive objection against the second sub-thesis. For the reluctance to regard the appearance of Koplik spots as explanatory may well reflect doubts as to whether, as a matter of universal law, those spots are always followed by the later manifestations of measles. Perhaps a local inoculation with a small amount of measles virus would produce the spots without leading to a full-blown case of the measles. If this were so, the appearance of the spots would still afford a usually reliable basis for predicting the occurrence of further symptoms, since exceptional conditions of the kind just mentioned would be extremely rare; but the generalization that Koplik spots are always followed by later symptoms of the measles would not express a law and thus could not properly support a corresponding D-N explanation.

The objection just considered concerns the explanatory potential of predictive arguments of the form (D-N). But the second sub-thesis, in its general form, which is not limited to D-N predictions, has further been challenged, particularly by Scheffler and by Scriven,[50] on the ground that there are other kinds of predictive argument that are adequate for scientific prediction, yet not for explanation. Specifically, as Scheffler notes, a scientific prediction may be based on a finite set of data which includes no laws and which would have no explanatory force. For example, a finite set of data obtained in an extensive test of the hypothesis that the electric resistance of metals increases with their temperature may afford good support for that hypothesis and may thus provide an acceptable basis for the prediction that in an as yet unexamined instance, a rise in temperature in a metal conductor will be accoumpanied by an increase in resistance. But if this event then actually occurs, the test data clearly do not provide an explanation for it. Similarly, a list of the results obtained in a long series of tossings of a given coin may provide a good basis for predicting the percentage of Heads and Tails to be expected in the next 1000 tossings of the same coin; but again, that list of data provides no explanation for the subsequent results. Cases like these raise the question of whether there are not sound modes of scientific prediction that proceed from particulars to particulars without benefit of general laws such as seem to be required for any adequate explanation. Now, the predictive arguments just considered are not deductive but probabilistic in character; and the role of probabilistic inference for explanation and prediction will be considered more fully in section 3 of this essay. But in regard to the second sub-thesis of the structural identity claim, let us note this much here: the predictions in our illustrations proceed from an observed sample of a population to another, as yet unobserved one; and on some current theories

50. See Scheffler (1957), p. 296 and (1963), p. 42; Scriven (1959a), p. 480.

of probabilistic inference such arguments do not depend upon the assumption of general empirical laws. According to Carnap's theory of inductive logic,[51] for example, such inferences are possible on purely logical grounds; the information about the given sample confers a definite logical probability upon any proposed prediction concerning an as yet unobserved sample. On the other hand, certain statistical theories of probabilistic inference eschew the notion of purely logical probabilities and qualify predictions of the kind here considered as sound only on the further assumption that the selection of individual cases from the total population has the character of a random experiment with certain general statistical characteristics. But that assumption, when explicitly spelled out, has the form of a general law of statistic-probabilistic form; hence, the predictions are effected by means of covering laws after all. And though these laws do not have the strictly universal character of those invoked in D-N explanations and predictions, they can serve in an explanatory capacity as well. Thus construed, even the predictions here under discussion turn out to be (incompletely formulated) potential explanations.

The basic questions at issue between these different conceptions of probabilistic inference are still the subject of debate and research, and this essay is not the place to attempt a fuller appraisal of the opposing views. The second sub-thesis of the structural identity claim for explanation and prediction will therefore be regarded here as an open question.

3. STATISTICAL EXPLANATION

3.1 LAWS OF STATISTICAL FORM. We now turn our attention to explanations based on nomological statements of a kind we have not so far considered, which have come to play an increasingly important role in empirical science. I will refer to them as *laws or theoretical principles of statistic-probabilistic form*, or as *statistical laws*, for short.

Most of our discussion will be concerned with the explanatory use of statistical laws of a very simple kind; we will call them *laws of basic statistical form*. These are statements to the effect that the statistical probability for an event of kind F to be also of kind G is r, or that

$$p(G,F) = r$$

for short. Broadly speaking, this statement asserts that in the long run the proportion of those instances of F which are also instances of G is approximately r. (A fuller account will be given in section 3.3.)

For example, the statement that the rolling of a given slightly irregular die

51. Carnap (1950), section 110.

(event of kind F) yields an ace (event of kind G) with a probability of .15, i.e., in about 15 per cent of all cases in the long run, has this basic statistical form. And so does the law that the half-life of radon is 3.82 days, i.e., that the statistical probability for a radon atom to disintegrate during any given period of 3.82 days is 1/2, which means, roughly, that of a sample of radon containing a large number of atoms, very close to one half of the atoms decay within 3.82 days.

Laws of basic statistical form may be regarded as less stringent counterparts of laws that have the universal conditional form

$$(x)\,(Fx \supset Gx)$$

asserting that any instance of F is an instance of G, as for example: 'Any gas expands when heated under constant pressure'. Indeed, the two kinds of law share an important feature, which is symptomatic of their nomological character: both make general claims concerning a class of cases that might be said to be potentially infinite. As we noted earlier, a statement which is logically equivalent to a finite conjunction of singular sentences, and which in this sense makes a claim concerning only a finite class of cases, does not qualify as a law and lacks the explanatory force of a nomological statement. Lawlike sentences, whether true or false, are not just conveniently telescoped summaries of finite sets of data concerning particular instances.

For example, the law that gases expand when heated under constant pressure is not tantamount to the statement that in all instances that have so far been observed, or perhaps in all instances that have so far occurred, an increase in the temperature of a gas under constant pressure has been accompanied by an increase in volume. Rather it asserts that a growth in volume is associated with the heating of a gas under constant pressure in *any* case, whether past, present, or future, and whether actually observed or not. It even implies counterfactual and subjunctive conditionals to the effect that if a given body of gas had been heated or were to be heated under constant pressure, its volume would have increased, or would increase, as well.

Similarly, the probabilistic laws of genetics or of radioactive decay are not tantamount to descriptive reports of the frequencies with which some kind of phenomenon has been found to occur in a finite class of observed cases: they assert certain peculiar, namely probabilistic, modes of connection between potentially infinite classes of occurrences. In a statistical law of basic form, as contradistinguished from a statistical description specifying relative frequencies in some finite set, the "reference class" F is not assumed to be finite. Indeed, we might say that a law of the form '$p\,(G,F) = r$' refers not only to all actual instances of F, but, so to speak, to the class of all its potential instances. Suppose, for example, that we are given a homogeneous regular tetrahedron whose faces are marked 'I', 'II', 'III', 'IV'. We might then assert that the probability of obtaining a III, i.e., of the tetrahedron's coming to rest on that face upon being tossed out of a dice box, is 1/4. But, while this assertion says something about the frequency with which a III is obtained as a result of rolling the tetrahedron, it cannot be construed as simply specifying that frequency for the class of all tosses which are, in fact, ever performed with the tetrahedron. For we might well maintain our hypothesis even if we were informed that the tetrahedron would actually be tossed only a few times throughout its existence, and in this case, our probability statement would surely not be meant to assert that exactly, or even nearly, one-fourth of those tosses would yield the result III. Moreover, our statement would be perfectly meaningful and might, indeed, be well supported (e.g., by results obtained with similar tetrahedra or with other homogeneous bodies in the form of regular solids) even if the given tetrahedron happened to be destroyed without ever having been tosed at all. What the probability statement attributes to the tetrahedron is, therefore, not the frequency with which the result III is obtained in actual past or future rollings, but a certain *disposition*, namely, the disposition to yield the result III in about one out of four cases, in the long run. This disposition might be characterized by means of a subjunctive conditional phrase: if the tetrahedron were to be tossed a large number of times, it would yield the result III in about one-fourth of the cases.[1] Implications in the form of counterfactual and subjective conditionals are thus hallmarks of lawlike statements both of strictly universal and of statistical form.

As for the distinction between lawlike sentences of strictly universal form and those of probabilistic or statistical form, it is sometimes thought that statements asserting strictly universal connections, such as Galileo's law or Newton's law of gravitation, rest, after all, only on a finite and thus inevitably incomplete body of evidence; that, therefore, they may well have as yet

1. Carnap (1951-54, pp. 190-92) has argued in a similar vein that the statistical probability of rolling an ace with a given die is a physical characteristic, which he also calls "the probability state" of the die, and that the relative frequency with which rollings of the die yield an ace is a symptom of that state, much as the expansion of the mercury column in a thermometer is a symptom of its temperature state.

The dispositional construal I have outlined for the concept of statistical probability appears to be in close accord also with the "propensity interpretation" advocated by Popper. The latter "differs from the purely statistical or frequency interpretation only in this—that it considers the probability as a characteristic property of the experimental arrangement rather than as a property of a sequence"; the property in question is explicitly construed as *dispositional*. (Popper 1957, pp. 67-68). See also the discussion of this paper in Körner (1957), pp. 78-89, *passim*. However, the currently available statements of the propensity interpretation are all rather brief; a fuller presentation is to be given in a forthcoming book by Popper.

undetected exceptions; and that accordingly they, too, should be qualified as only probabilistic. But this argument confounds the claim made by a given statement with the evidence available in support of it. On the latter score, all empirical statements are only more or less well supported by the relevant evidence at our disposal; or, in the parlance of some theorists, they have a more or less high logical or inductive probability conferred upon them by that evidence. But the distinction between lawlike statements of strictly universal form and those of probabilistic form pertains, not to the evidential support of the statements in question, but to the claims made by them: roughly speaking, the former attribute (truly or falsely) a certain characteristic to all members of a certain class; the latter, to a specified proportion of its members.

Even if all the supposedly universal laws of empirical science should eventually come to be regarded as reflections of underlying statistical uniformities—an interpretation that the kinetic theory of matter gives to the classical laws of thermodynamics, for example—even then the distinction between the two types of law and the corresponding explanations is not wiped out: in fact, it is presupposed in the very formulation of the conjecture.

Nor is a statement of the universal conditional form

$$(x)(Fx \supset Gx)$$

logically equivalent to the corresponding statement of the basic statistical form

$$p(G, F) = 1$$

for, as will be shown more fully in section 3.3, the latter asserts only that it is practically certain that in a large number of instances of F, almost all are instances of G; hence the probability statement may be true even if the corresponding statement of strictly universal form is false.

So far, we have dealt only with statistical laws of basic form. Let us now say more generally that *a statement has the form of a statistical law*, or is of probabilistic-statistical character, if it is formulated in terms of statistical probabilities, i.e., if it contains (nonvacuously) the term 'statistical probability' or some notational equivalent, or a term—such as 'half-life'—which is defined by means of statistical probabilities.

Take, for example, the statement that when two coins are flipped simultaneously, the face shown by one is independent of that shown by the other. This amounts to saying that the probability for the second coin to show heads when the first shows heads is the same as when the first shows tails; and vice versa. Generally, assertions of statistical independence have the form of statistical laws, though they are not of basic statistical form. Similarly, a statement asserting a statistical dependence or "aftereffect" has the form of a statistical law; for example, the statement that in any given area the probability for a day to be cloudy when it follows a cloudy day is greater than when it follows a noncloudy day. Still other laws of statistical form are formulated in terms of mean values of certain variables, such as the mean kinetic energy and the mean free path of the molecules in a gas; the notion of mean value here invoked is defined by reference to statistical probabilities.

By a *statistical explanation*, let us now understand any explanation that makes essential use of at least one law or theoretical principle of statistical form. In the following subsections, we will examine the logical structure of such explanations. We will find that there are two logically different types of statistical explanation. One of them amounts, basically, to the deductive subsumption of a narrower statistical uniformity under more comprehensive ones: I will call it *deductive-statistical explanation*. The other involves the subsumption, in a peculiar nondeductive sense, of a particular occurrence under statistical laws; for reasons to be given later, it will be called *inductive-statistical explanation*.

3.2 DEDUCTIVE-STATISTICAL EXPLANATION. It is an instance of the so-called gambler's fallacy to assume that when several successive tossings of a fair coin have yielded heads, the next toss will more probably yield tails than heads. Why this is not the case can be explained by means of two hypotheses that have the form of statistical laws. The first is that the random experiment of flipping a fair coin yields heads with a statistical probability of 1/2. The second hypothesis is that the outcomes of different tossings of the coin are statistically independent, so that the probability of any specified sequence of outcomes—such as heads twice, then tails, then heads, then tails three times—equals the product of the probabilities of the constituent single outcomes. These two hypotheses in terms of statistical probabilities imply *deductively* that the probabliity for heads to come up after a long sequence of heads is still 1/2.

Certain statistical explanations offered in science are of the same deductive character, though often quite complex mathematically. Consider, for example, the hypothesis that for the atoms of every radioactive substance there is a characteristic probability of disintegrating during a given unit time internal, and that probability is independent of the age of the atom and of all external circumstances. This complex statistical hypothesis explains, by deductive implication, various other statistical aspects of radioactive decay, among them, the following: Suppose that the decay of individual atoms of some radioactive substance is recorded by means of the scintillations produced upon a sensitive screen by the alpha particles emitted by the disintegrating atoms. Then the time intervals separating successive scintillations will vary considerably in length, but intervals of different lengths will occur with different statistical probabilities. Specifically, if the mean time interval between successive scintillations is s seconds, then the probability for two successive scintillations to be separated by more than $n \cdot s$ seconds is $(1/e)^n$, where e is the base of the natural logarithms.[2]

Explanations of the kind here illustrated will be called *deductive-statistical explanations*, or *D-S explanations*. They involve the deduction of a statement in the form of a statistical law from an explanans that contains indispensably at least one law or theoretical principle of statistical form. The deduction is effected by means of the mathematical theory of statistical probability, which makes it possible to calculate certain derivative probabilities (those referred to in the explanandum) on the basis of other probabilities (specified in the explanans) which have been empirically ascertained or hypothetically assumed. What a D-S explanation accounts for is thus always a general uniformity expressed by a presumptive law of statistical form.

Ultimately, however, statistical laws are meant to be applied to particular occurrences and to establish explanatory and predictive connections among them. In the next subsection, we will examine the statistical explanation of particular events. Our discussion will be limited to the case where the explanatory statistical laws are of basic form: this will suffice to exhibit the basic logical differences between the statistical and the deductive-nomological explanation of individual occurrences.

3.3 Inductive-Statistical Explanation. As an explanation of why patient John Jones recovered from a streptococcus infection, we might be told that Jones had been given penicillin. But if we try to amplify this explanatory claim by indicating a general connection between penicillin treatment and the subsiding of a streptococcus infection we cannot justifiably invoke a general law to the effect that in all cases of such infection, administration of penicillin will lead to recovery. What can be asserted, and what surely is taken for granted here, is only that penicillin will effect a cure in a high percentage of cases, or with a high statistical probability. This statement has the general character of a law of statistical form, and while the probability value is not specified, the statement indicates that it is high. But in contrast to the cases of deductive-nomological and deductive-statistical explanation, the explanans consisting of this statistical law together with the statement that the patient did receive penicillin obviously does not imply the explanandum statement, 'the patient

2. *Cf.* Mises (1939), pp. 272-78, where both the empirical findings and the explanatory argument are presented. This book also contains many other illustrations of what is here called deductive-statistical explanation.

recovered', with deductive certainty, but only, as we might say, with high likelihood, or near-certainty. Briefly, then, the explanation amounts to this argument:

(3a) The particular case of illness of John Jones—let us call it *j*—was an instance of severe streptococcal infection (*Sj*) which was treated with large doses of penicillin (*Pj*); and the statistical probability $p\ (R, S{\cdot}P)$ of recovery in cases where *S* and *P* are present is close to 1; hence, the case was practically certain to end in recovery (*Rj*).

This argument might invite the following schematization:

$$\text{(3b)}\qquad \frac{\begin{array}{l} p\ (R, S{\cdot}P) \text{ is close to } 1 \\ Sj \cdot Pj \end{array}}{\text{(Therefore:) It is practically certain (very likely) that } Rj}$$

In the literature on inductive inference, arguments thus based on statistical hypotheses have often been construed as having this form or a similar one. On this construal, the conclusion characteristically contains a modal qualifier such as 'almost certainly', 'with high probability', 'very likely', etc. But the conception of arguments having this character is untenable. For phrases of the form 'it is practically certain that *p*' or 'It is very likely that *p*', where the place of '*p*' is taken by some statement, are not complete self-contained sentences that can be qualified as either true or false. The statement that takes the place of '*p*'—for example, '*Rj*'—is either true or false, quite independently of whatever relevant evidence may be available, but it can be qualified as more or less likely, probable, certain, or the like only *relative to some body of evidence*. One and the same statement, such as '*Rj*', will be certain, very likely, not very likely, highly unlikely, and so forth, depending upon what evidence is considered. The phrase 'it is almost certain that *Rj*' taken by itself is therefore neither true nor false; and it cannot be inferred from the premises specified in (3b) nor from any other statements.

The confusion underlying the schematization (3b) might be further illuminated by considering its analogue for the case of deductive arguments. The force of a deductive inference, such as that from 'all *F* are *G*' and '*a* is *F*' to '*a* is *G*', is sometimes indicated by saying that if the premises are true, then the conclusion is necessarily true or is certain to be true—a phrasing that might suggest the schematization

$$\frac{\begin{array}{l} \text{All } F \text{ are } G \\ a \text{ is } F \end{array}}{\text{(Therefore:) It is necessary (certain) that } a \text{ is } G.}$$

But clearly the given premises—which might be, for example, 'all men are mortal' and 'Socrates is a man'—do not establish the sentence 'a is G' ('Socrates is mortal') as a necessary or certain turth. The certainty referred to in the informal paraphrase of the argument is relational: the statement 'a is G' is certain, or necessary, *relative to the specified premises*; i.e., their truth will guarantee its truth—which means nothing more than that 'a is G' is a logical consequence of those premises.

Analogously, to present our statistical explanation in the manner of schema (3b) is to misconstrue the function of the words 'almost certain' or 'very likely' as they occur in the formal wording of the explanation. Those words clearly must be taken to indicate that on the evidence provided by the explanans, or relative to that evidence, the explanandum is practically certain or very likely, i.e., that

(3c) 'Rj' is practically certain (very likely) relative to the explanans containing the sentences '$p\ (R, S{\cdot}P)$ is close to 1' and '$Sj \cdot Pj$'.[3]

The explanatory argument misrepresented by (3b) might therefore suitably be schematized as follows:

$$\text{(3d)}\qquad \begin{array}{l} p\,(R, S{\cdot}P) \text{ is close to } 1 \\ Sj \cdot Pj \\ \hline\hline Rj \end{array} \quad \text{[makes practically certain (very likely)]}$$

In this schema, the double line separating the "premises" from the "conclusion" is to signify that the relation of the former to the latter is not that of deductive implication but that of inductive support, the strength of which is indicated in square brackets.[4,5]

3. Phrases such as 'It is almost certain (very likely) that j recovers', even when given the relational construal here suggested, are ostensibly concerned with relations between propositions, such as those expressed by the sentences forming the conclusion and the premises of an argument. For the purpose of the present discussion, however, involvement with propositions can be avoided by construing the phrases in question as expressing logical relations between corresponding *sentences*, e.g., the conclusion-sentence and the premise-sentence of an argument. This construal, which underlies the formulation of (3c), will be adopted in this essay, though for the sake of convenience we may occasionally use a paraphrase.

4. In the familiar schematization of deductive arguments, with a single line separating the premises from the conclusion, no explicit distinction is made between a weaker and a stronger claim, either of which might be intended; namely (i) that the premises logically imply the conclusion and (ii) that, in addition, the premises are true. In the case of our probabilistic argument, (3c) expresses a weaker claim, analogous to (i), whereas (3d) may be taken to express a "proffered explanation" (the term is borrowed from Scheffler, (1957), section 1) in which, in addition, the explanatory premises are—however tentatively—asserted as true.

5. The considerations here outlined concerning the use of terms like 'probably' and 'certainly' as modal qualifiers of individual statements seem to me to militate also against

(continued overleaf)

Our schematization thus reflects explicitly the understanding that 'almost certain', 'very likely', 'practically impossible' and similar expressions often used in the phrasing of probabilistic arguments, including explanations, do not stand for properties possessed by certain propositions or the corresponding sentences, but for relations that some sentences bear to others. According to this understanding, the notion of the explanans of (3d) making the explanandum almost certain or very likely is but a special case of the idea of a given statement or set of statements—let us call it the grounds or the evidence *e*—conferring more or less strong inductive support or confirmation or credibility upon some statement *h*. To clarify and systematically to elaborate the idea here sketchily characterized is, of course, the objective of various theories of inductive reason-

the notion of categorical probability statement that C. I. Lewis sets forth in the following passage (italics the author's):

> Just as 'If *D* then (certainly) *P*, and *D* is the fact,' leads to the categorical consequence, 'Therefore (certainly) *P*'; so too, 'If *D* then probably *P*, and *D* is the fact', leads to a categorical consequence expressed by 'It is probable that *P*'. And this conclusion is not merely the statement over again of the probability relation between '*P*' and '*D*'; any more than 'Therefore (certainly) *P*' is the statement over again of 'If *D* then (certainly) *P*'. 'If the barometer is high, tomorrow will probably be fair; and the barometer *is* high', categorically assures something expressed by 'Tomorrow will probably be fair'. This probability is still relative to the grounds of judgment; but if these grounds are actual, and contain all the available evidence which is pertinent, then it is not only categorical but may fairly be called *the* probability of the event in question. (1946, p. 319).

This position seems to me to be open to just those objections suggested in the main text. If '*P*' is a statement, then the expressions 'certainly P' and 'probably *P*' as envisaged in the quoted passage are not statements. If we ask how one would go about trying to ascertain whether they were true, we realize that we are entirely at a loss unless and until a reference set of statements or assumptions has been specified relative to which *P* may then be found to be certain, or to be highly probable, or neither. The expressions in question, then, are essentially incomplete; they are elliptic formulations of relational statements; neither of them can be the conclusion of an inference. However plausible Lewis's suggestion may seem, there is no analogue in inductive logic to *modus ponens*, or the "rule of detachment," of deductive logic, which, given the information that '*D*', and also 'if *D* then *P*', are true statements, authorizes us to detach the consequent '*P*' in the conditional premise and to assert it as a self-contained statement which must then be true as well.

At the end of the quoted passage, Lewis suggests the important idea that 'probably *P*' might be taken to mean that the total relevant evidence available at the time confers high probability upon *P*. But even this statement is relational in that it tacitly refers to some unspecified time, and, besides, his general notion of a categorical probability statement as a conclusion of an argument is not made dependent on the assumption that the premises of the argument include all the relevant evidence available.

It must be stressed, however, that elsewhere in his discussion, Lewis emphasizes the relativity of (logical) probability, and, thus, the very characteristic that rules out the conception of categorical probability statements.

Similar objections apply, I think, to Toulmin's construal of probabilistic arguments; *cf.* Toulmin (1958) and the discussion in Hempel (1960), sections 1-3.

ing. It is still a matter of debate to what extent clear criteria and a precise theory for the concept at issue can be developed. Several attempts have been made to formulate rigorous logical theories for a concept of inductive support that admits of numerical or nonnumerical gradations in strength: two outstanding examples of such efforts are Keynes's theory of probability and, especially, Carnap's impressive system of inductive logic.[6] In the latter, the degree to which a sentence, or hypothesis, h is confirmed by an evidence sentence e is represented by a function $c(h,e)$, whose values lie in the interval from 0 to 1 inclusive, and which satisfies all the basic principles of abstract probability theory; $c(h,e)$ is therefore also referred to as the *logical or inductive* probability of h on e. This concept of inductive probability as a quantitative logical relation between statements must be sharply distinguished from the concept of statistical probability as a quantitative empirical relation between kinds or classes of events. The two concepts have a common formal structure, however, in virtue of which both of them qualify as probabilities: both are defined, in their respective formal theories, in terms of nonnegative additive set functions whose values range from 0 to 1. Carnap's theory provides an explicit definition of $c(h,e)$ for the case where the sentences h and e belong to one or another of certain relatively simple kinds of formalized language; the extension of his approach to languages whose logical apparatus would be adequate for the formulation of advanced scientific theories is as yet an open problem.

But, independently of the extent to which the relation of the explanandum to the explanans can be analyzed in terms of Carnap's quantitative concept of inductive probability, probabilistic explanations must be viewed as inductive in the broad sense here adumbrated. To refer to the general notion of inductive support as capable of gradations, without commitment to any one particular theory of inductive support or confirmation, we will use the phrase '*(degree of) inductive support of h relative to e*'.[7]

Explanations of particular facts or events by means of statistic-probabilistic laws thus present themselves as arguments that are *inductive* or *probabilistic* in the sense that the explanans confers upon the explanandum a more or less high degree of inductive support or of logical (inductive) probability; they

6. See Keynes (1921); of Carnap's numerous writings on the subject, *cf.* especially (1945), (1950), (1952), (1962).

7. Some recent attempts to give precise explications of this general notion have led to concepts that do not have all the formal characteristics of a probability function. One such construal is presented in Helmer and Oppenheim (1945) and, less technically, in Hempel and Oppenheim (1945). Another is the concept of degree of factual support propounded and theoretically developed in Kemeny and Oppenheim (1952). For a suggestive distinction and comparison of different concepts of evidence, see Rescher (1958).

will therefore be called *inductive-statistical explanations*, or *I-S explanations*. Explanations, such as (3d), in which the statistical laws invoked are of basic form, will also be called *I-S explanations of basic form*.

I will now try to show that the inductive construal here suggested for the statistical explanation of particular facts is called for also by the empirical interpretation that probabilistic laws have received in recent versions of the theory of statistical probability and its applications.

The mathematical theory of statistical probability is intended to provide a theoretical account of the statistical aspects of repeatable processes of a certain kind, which are referred to as random processes or random experiments. Roughly, a random experiment is a kind of process or event which can be repeated indefinitely by man or by nature, and which yields in each case one out of a certain finite or infinite set of "results" or "outcomes" in such a way that while the outcomes vary from case to case in an irregular and unpredictable manner, the relative frequencies with which the different outcomes occur tend to become more or less constant as the number of performances increases. The flipping of a coin, with heads and tails as the possible outcomes, is a familiar example of a random experiment.

The theory of probability offers a "mathematical model" of the general mathematical properties and interrelations of the long-run frequencies associated with the outcomes of random experiments.

In the model, each of the different "possible outcomes" assigned to a given random experiment F is represented by a set G, which may be thought of as the set of those performances of the experiment that yield the outcome in question, while F may be viewed as the set of all performances of the random experiment. The probability of obtaining an outcome of a given kind G as a result of performing an experiment of kind F is then represented as a measure, $p_F(G)$, of the size of set G in relation to set F.

The postulates of the mathematical theory specify that p_F is a nonnegative additive set function whose maximum value is 1, i.e., for every possible outcome G of F, $p_F(G) \geqslant 0$; if G_1, G_2 are mutually exclusive outcomes of F, then $p_F(G_1 \vee G_2) = p_F(G_1) + p_F(G_2)$; and $p_F(F) = 1$. These stipulations permit the proof of the theorems of elementary probability theory; to deal with experiments that admit of infinitely many different outcomes, the requirement of additivity is suitably extended to infinite sequences of mutually exclusive outcome sets $G_1, G_2, G_3, \ldots$.

The resulting abstract theory is applied to empirical subject matter by means of an interpretation that relates statements in terms of probabilities as set-measures to statements about long-run relative frequencies associated with the outcomes of random experiments. I will now state this interpretation in a formulation which is essentially that given by Cramér.[8] For convenience, the notation '$P_F(G)$' will henceforth be replaced by '$p(G,F)$'.

(3e) *Frequency interpretation of statistical probability.* Let F be a given kind of random experiment and G a possible result of it; then the statement that $p(G, F)=r$ means that in a long series of repetitions of F, it is practically certain that the relative frequency of the result G will be approximately equal to r.

Cramér also states two corrollaries of this interpretation which refer to those cases where r differs very little from 0 or from 1; they are of special interest for our further discussion of probabilistic explanation. I will therefore note them here, again following Cramér's formulation in its essentials.[9]

(3e.1) If $1 - p(G,F) < \varepsilon$, where ε is some very small positive number, then if random experiment F is performed one single time, it is practically certain that the result G will occur.

(3e.2) If $p(G,F)<\varepsilon$, where ε is some very small positive number, then if random experiment F is performed one single time, it is practically certain that result G will not occur.

As the frequency interpretation here formulated makes use of such vague phrases as 'a long series', 'practically certain', 'approximately equal', and the like, it clearly does not provide a precise definition of statistical probabilities in terms of observable relative frequencies. But some vagueness appears to be inevitable if the mathematical calculus of probability is to serve as a theoretical representation of the mathematical relations among empirically ascertained relative frequencies which remain only approximately constant when the observed sample increases.[10]

8. See Cramér (1946), pp. 148-49. Cramér's book includes a detailed discussion of the foundations of statistical probability theory and its applications. Similar formulations of the frequency interpretation have been given by earlier representatives of this measure-theoretical conception of statistical probability; for example, by Kolmogoroff (1933, p. 4).

9. For (3e.1), see Cramér (1946), p. 150; for (3e.2), see Cramér (1946), p. 149 and the very similar formulation in Kolmogoroff (1933), p. 4.

10. In certain forms of the mathematical theory, the statistical probability of a given outcome is explicitly defined, namely, as the limit of the relative frequency of that outcome in an infinite series of performances of the pertinent random experiment. Two important variants of this approach were developed by Mises, *cf.* (1931), (1939) and by Reichenbach, *cf.* (1949). But infinite series of performances are not realizable or observable, and the limit-definition of statistical probability thus provides no criteria for the application of that concept to observable empirical subject matter. In this respect the limit-construal of probability is an idealized theoretical concept, and criteria for its empirical application will again have to involve some vague terms of the kind resorted to in (3c) and its corollaries. In particular, a statement specifying the limit of the relative frequency of the result G in an infinite sequence of performances of random experiment F has no deductive implications concerning the frequency of G in any finite set of performances, however large it may be. The relation between

(continued overleaf)

Of particular interest for an analysis of I-S explanation, however, is the fact that the phrase 'it is practically certain that' occurs in the general statement (3e) of the statistical interpretation and that its two special corollaries (3e.1) and (3e.2) still contain that phrase, though they manage to avoid the vague expressions 'a long series of repetitions' and 'approximately equal'. The function of the words 'it is practically certain that' is clear: they indicate that the logical connection between statistical probability statements and the empirical frequency statements associated with them is inductive rather than deductive. This point can be made more explicit by restating (3e) as follows: The information that $p(G,F)=r$ and that S is a set of n performances of F, where n is a large number, confers near-certainty (high inductive support) upon the statement that the number of those performances in S whose outcome is G is approximately $n \cdot r$. The two corollaries admit of an analogous construal. Thus, (3e.1) may be restated as follows: The information that $1\text{-}p(G,F) < \varepsilon$ (where ε is a small positive number) and that individual event i is a performance of random experiment F (or that Fi, for short) lends strong inductive support to the statement that i yields outcome G, or that Gi, for short. Or, in a slightly different phrasing: 'Gi' is practically certain relative to the two sentences '$p(G,F)$ is very close to 1' and 'Fi'. This last version has the same form as (3c); thus, in giving an inductive construal to the explanatory import of probabilistic laws in the manner illustrated by (3d), we are in basic accord with the empirical interpretation given to probabilistic laws in the contemporary theory of statistical probability.[11]

In our example concerning recovery from a streptococcus infection, the statistical law invoked did not specify a definite numerical value for the probability of effecting recovery by means of penicillin. Now we will consider a simple case of I-S explanation in which the relevant probability statement is quite specific. Let the experiment D (more exactly, an experiment of *kind D*) consist in drawing, with subsequent replacement, a ball from an urn containing 999 white balls and one black, all of the same size and material. We might then accept the statistical hypothesis that with respect to the outcomes "white ball"

11. However, the representatives of current statistical probability theory do not, in general, take explicit notice of the inductive character of their statistical interpretation of probability statements. Even less do they attempt to analyze the inductive concept of practical certainty, which clearly falls outside the mathematical theory that is their principal concern.

---

probability statements thus construed and the corresponding statements about relative frequencies in finite runs must therefore again be viewed as inductive.

For a concise account of the limit conception of statistical probability and a lucid discussion of some of its difficulties, see Nagel (1939), especially sections 4 and 7.

and "black ball," D is a random experiment in which the probability of obtaining a white ball is $p(W,D) = .999$. According to the statistical interpretation, this is a hypothesis susceptible of test by reference to finite statistical samples, but for our present purposes, we need not consider the grounds we might have for accepting the hypothesis; for we are concerned only with its explanatory use. Our rule (3e.1) suggests that the hypothesis might indeed be used to explain probabilistically the results of certain individual drawings from the urn, i.e., the results of certain performances of D. Suppose, for example, that a particular drawing, d, produces a white ball. Since $p(W,D)$ differs from 1 by less than, say, .0011, which is quite a small amount, rule (3e.1) suggests the following explanatory argument in analogy to (3d):

$$\text{(3f)} \qquad \begin{array}{c} 1 - p(W,D) < .0011 \\ Dd \\ \hline\hline Wd \end{array} \text{ [makes practically certain]}$$

Again, the explanans here does not logically imply the explanandum; and the argument does not show that, assuming the truth of the statements adduced in the explanans, the explanandum phenomenon was to be expected "with certainty." Rather, the argument may be said to show that on the information provided by the explanans, the explanandum event was to be expected with "practical" certainty, or with very high likelihood.

Carnap's conception of inductive logic suggests that the vague phrase 'makes practically certain', which appears between brackets in (3f), might be replaced by a more definite quantitative one. This would call for an extension of Carnap's theory to languages in which statistical probability statements can be formulated. While the logical apparatus of the languages covered by Carnap's published work is not rich enough for this purpose,[12] it seems clear that in cases of the simple kind exemplified by (3f), the numerical value of the logical probability should equal that of the corresponding statistical probability. For example, the information that with statistical probability .999, a drawing from the urn will produce a white ball, and that the particular event d is a drawing from the urn, should confer a logical probability of .999 upon the "conclusion" that the ball produced by d is white. More generally, this rule may be stated as follows:

(3.g) If e is the statement '$(p(G,F)=r) \cdot F\,b$', and h is '$G\,b$', then $c(h, e) = r$.

This rule is in keeping with the conception, set forth by Carnap, of logical probability as a fair betting quotient for a bet on h on the basis of e. It accords

12. According to a personal communication from Professor Carnap, his system has by now been extended in that direction.

equally with Carnap's view that the logical probability on evidence *e* of the hypothesis that a particular case *b* will have a specified property *M* may be regarded as an estimate, based on *e*, of the relative frequency of *M* in any class *K* of cases on which the evidence *e* does not report. Indeed, Carnap adds that the logical probability of '*Mb*' on *e* may in certain cases be considered as an estimate of the statistical probability of *M*.[13] If, therefore, *e* actually contains the information that the statistical probability of *M* is *r*, then the estimate, on *e*, of that statistical probability, and thus of the logical probability of '*Mb*' on *e*, should clearly be *r* as well.

And just as the rule (3e.1) provides the logical rationale for statistical explanations such as (3f), so our rule (3g) provides the rationale for a similar kind of probabilistic explanation, which invokes quantitatively definite statistical laws and which may be schematized as follows:

$$\text{(3h)}\qquad \begin{array}{l} p(G,F) = r \\ F\,i \\ \hline\hline G\,i \end{array} \quad [r]$$

An explanatory argument of this form would serve to account for the fact that a given individual case *i* exhibits the characteristic *G* by pointing out that *i* is a case of *F*; that the statistical probability for an *F* to exhibit characteristic *G* is *r*; and that, according to rule (3g), this explanatory information confers the logical probability *r* upon the explanandum statement. I will refer to *r* also as the probability *associated with* the explanation. Of course, an argument of this kind will count as explanatory only if the number *r* is fairly close to 1. But it seems impossible, without being arbitrary, to designate any particular number, say .8, as the minimum value of the probability *r* permissible in an explanation.

In our example, the probabilistic explanation of the drawing of a white ball may now be put into the form (3h) as follows:

$$\text{(3i)}\qquad \begin{array}{l} p(W,D) = .999 \\ D\,d \\ \hline\hline W\,d \end{array} \quad [.999]$$

Now, it is often said that probabilistic laws can serve to account for statistical aspects of large samples, but surely can explain nothing about an individual case. Examples like the following might seem to bear out this contention. The law that the flipping of a regular coin yields heads with the probability 1/2

13. Carnap (1950), pp. 168-75.

clearly does not enable us to explain why a particular flipping produced heads; whereas the same law (plus the assumption that the results of different flippings are statistically independent of each other) may be used to account for the fact that the number of heads obtained in a particular series of 10,000 flippings fell between 4,9000 and 5,100; for this outcome has a probability exceeding .95. But if we count this outcome as explained because of the high probability the explanans confers upon it, then clearly we must also grant explanatory status to arguments such as (3i) whose explanans makes it highly probable that the given outcome will occur if the relevant random experiment is performed just once.

It is also sometimes thought that because probabilistic arguments are not logically conclusive they cannot serve to explain; for even if the explanans is true, it is still possible that the explanandum phenomenon might not have come about;[14] in the case of (3i), for example, drawing *d* might have produced a black ball despite the high probability for a white one to be drawn. But this objection to the idea of probabilistic explanation rests on a too restrictive conception of scientific explanation; for many important explanatory accounts offered by empirical science make quite explicit use of statistical laws which, in conjunction with the rest of the explanatory information adduced, make the explanandum no more than highly probable.

For example, by means of Mendelian genetic principles it can be shown to be highly probable that in a random sample taken from a population of pea plants each of whose parent plants represents a cross of a pure white-flowered and a pure red-flowered strain, approximately 75 per cent of the plants will have red flowers and the rest, white ones. This argument, which may be used for explanatory or for predictive purposes, is inductive-statistical; what it explains or predicts are the approximate percentages of red- and white-flowered plants in the sample. The "premises" by reference to which the specified percentages are shown to be highly probable include (1) the pertinent laws of genetics, some of which have statistical, others strictly universal form; and (2) information of the kind mentioned above about the genetic make-up of the parent generation of the plants from which the sample is taken. The genetic principles of strictly universal form include the laws that the colors in question are tied to specific genes, that the red gene is dominant over the white one,

14. Thus Scriven (1959, p. 467), says that "statistical statements are too weak—they abandon the hold on the individual case. . . . An event can rattle around inside a network of statistical laws." Dray (1963, p. 119), expresses a similar view. These observations are quite correct if they are simply meant to say that statistical laws have no deductive implications concerning particular events, but they are misleading if they are used to suggest that statistical laws can have no explanatory significance for particular occurrences.

and various other general laws concerning the transmission, by genes, of the colors in question—or, perhaps, of a broader set of gene-linked traits. Among the statistical generalizations invoked is the hypothesis that the four possible combinations of color-determining genes—WW, WR, RW, RR—are statistically equiprobable in their occurrence in the offspring of two plants of the hybrid generation.

Let us now examine somewhat more closely an explanatory use of the law for radioactive decay of radon, which states that this element has a half-life of 3.82 days. This law may be invoked for a statistical explanation of the fact that within 7.64 days, a particular sample consisting of 10 milligrams of radon was reduced, by radioactive decay, to a residual amount falling somewhere within the interval from 2.4 to 2.6 milligrams; it could similarly be used for predicting a particular outcome of this kind. The gist of the explanatory and predictive arguments is this: The statement giving the half-life of radon conveys two statistical laws, (i) the statistical probability for an atom of radon to undergo radioactive decay within a period of 3.82 days is 1/2, and (ii) the decay of different radon atoms constitutes statistically independent events. One further premise used is the statement that the number of atoms in 10 milligrams of radon is enormously large (in excess of 10^{19}). As mathematical probability theory shows, the two laws in conjunction with this last statement imply deductively that the statistical probability is exceedingly high that the mass of the radon atoms surviving after 7.64 days will not deviate from 2.5 milligrams by more than .1 milligrams, i.e., that it will fall within the specified interval. More explicitly, the consequence deducible from the two statistical laws in conjunction with the information on the large number of atoms involved is another statistical law to this effect: The statistical probability is very high that the random experiment F of letting 10 milligrams of radon decay for 7.68 days will yield an outcome of kind G, namely a residual amount of radon whose mass falls within the interval from 2.4 to 2.6 milligrams. Indeed, the probability is so high that, according to the interpretation (9.2b), if the experiment F is performed just one single time, it is "practically certain" that the outcome will be of kind G. In this sense, it is rational on the basis of the given information to expect the outcome G to occur as the result of a single performance of F. Also in this sense, the information concerning the half-life of radon and the large number of atoms involved in an experiment of kind F affords a statistical explanation or prediction of the occurrence of G in a particular performance of the experiment.

By way of another illustration, take the problem of explaining certain quantitative aspects of the Brownian movement displayed by small particles suspended in a liquid—a phenomenon qualitatively explained as resulting from the irregular impacts, upon the suspended particles, of the surrounding molecules in thermal agitation. From assumptions based on the probabilistic principles of the kinetic theory of heat, Einstein derived a law to the effect that the mean displacement of such particles is proportional to the square root of the elapsed time.[15] But the theoretical definition of the mean displacement is formulated in terms of the statistical probabilities of the various possible displacements, and Einstein's law is therefore probabilistic in character. Hence it does not logically imply definite values for the average displacement exhibited by finite numbers of particles. But the law makes it highly probable, in the sense discussed above, that the average displacements in finite samples will be very nearly proportional to the square root of the elapsed time—and this has indeed been found to be the case. Thus, Einstein's law provides a probabilistic explanation for observed aspects of Brownian movement.

As is illustrated by these examples and by others that will be considered soon, accounts in terms of statistical laws or theories thus play a very important role in science. Rather than deny them explanatory status on the ground that nonrealization of the explanandum is compatible with the explanans, we have to acknowledge that they constitute explanations of a distinct logical character, reflecting, we might say, a different sense of the word 'because'. Mises expresses this point of view when, contemplating recent changes in the notion of causality, he anticipates that "people will gradually come to be satisfied by causal statements of this kind: It is *because* the die was loaded that the 'six' shows more frequently (but we do not know what the next number will be); or: *Because* the vacuum was heightened and the voltage increased, the radiation became more intense (but we do not know the precise number of scintillations that will occur in the next minute)."[16] This passage clearly refers to statistical explanation in the sense here under consideration; it sets forth what might be called a statistical-probabilistic concept of "because," in contradistinction to a strictly deterministic one, which would correspond to deductive-nomological explanation.

Our discussion of the statistical explanation of particular occurrences has so far been concerned to exhibit its inductive character. In the next subsection, we will consider a further important characteristic which sets I-S explanation sharply apart from its deductive counterparts.

[illegible] For details, and for a full account of some experimental tests of this formula, see Svedber[illegible]12), pp. 89 ff. The basic ideas of the probabilistic explanation of some other quantitati[illegible]ects of Brownian movement are lucidly presented in Mises (1939), pp. 259-68.

16. Mises (1951), p. 188, italics the author's.

3.4 The Ambiguity of inductive-Statistical Explanation and the Requirement of Maximal Specificity.

3.4.1. *The Problem of Explanatory Ambiguity*. Consider once more the explanation (3d) of recovery in the particular case j of John Jones's illness. The statistical law there invoked claims recovery in response to penicillin only for a high percentage of streptococcal infections, but not for all of them; and in fact, certain streptococcus strains are resistant to penicillin. Let us say that an occurrence, e.g., a particular case of illness, has the property S^* (or belongs to the class S^*) if it is an instance of infection with a penicillin-resistant streptococcus strain. Then the probability of recovery among randomly chosen instances of S^* which are treated with penicillin will be quite small, i.e., $p(R, S^* \cdot P)$ will be close to 0 and the probability of non-recovery, $p(\overline{R}, S^* \cdot P)$ will be close to 1. But suppose now that Jones's illness is in fact a streptococcal infection of the penicillin-resistant variety, and consider the following argument:

$$\text{(3k)} \qquad \frac{\begin{array}{l} p(\overline{R}, S^* \cdot P) \text{ is close to } 1 \\ S^*j \,.\, Pj \end{array}}{\overline{R}j} \qquad \text{[makes practically certain]}$$

This "rival" argument has the same form as (3d), and on our assumptions, its premises are true, just like those of (3d). Yet its conclusion is the contradictory of the conclusion of (3d).

Or suppose that Jones is an octogenarian with a weak heart, and that in this group, S^{**}, the probability of recovery from a streptococcus infection in response to penicillin treatment, $p(R, S^{**} \cdot P)$, is quite small. Then, there is the following rival argument to (3d), which presents Jones's nonrecovery as practically certain in the light of premises which are true:

$$\text{(3l)} \qquad \frac{\begin{array}{l} p(\overline{R}, S^{**} \cdot P) \text{ is close to } 1 \\ S^{**}j \cdot Pj \end{array}}{\overline{R}j} \qquad \text{[makes practically certain]}$$

The peculiar logical phenomenon here illustrated will be called the *ambiguity of inductive-statistical explanation* or, briefly, of *statistical explanation*. This ambiguity derives from the fact that a given individual event (e.g., Jones's illness) will often be obtainable by random selection from any one of several "reference classes" (such as $S \cdot P$, $S^* \cdot P$, $S^{**} \cdot P$), with respect to which the kind of occurrence (e.g., R) instantiated by the given event has very different statistical probabilities. Hence, for a proposed probabilistic explanation with true explanans which confers near-certainty upon a particular event, there will often exist a rival argument of the same probabilistic form and with equally true premises which confers near-certainty upon the nonoccurrence of the same event. And any statistical explanation for the occurrence of an event must seem suspect if there is the possibility of a logically and empirically equally sound probabilistic account for its nonoccurrence. *This predicament has no analogue in the case of deductive explanation*; for if the premises of a proposed deductive explanation are true then so is its conclusion; and its contradictory, being false, cannot be a logical consequence of a rival set of premises that are equally true.

Here is another example of the ambiguity of I-S explanation: Upon expressing surprise at finding the weather in Stanford warm and sunny on a date as autumnal as November 27, I might be told, by way of explanation, that this was rather to be expected because the probability of warm and sunny weather (W) on a November day in Stanford (N) is, say, .95. Schematically, this account would take the following form, where 'n' stands for 'November 27':

$$\text{(3m)} \qquad \begin{array}{l} p\,(W,N) = .95 \\ Nn \\ \hline\hline Wn \end{array} \quad [.95]$$

But suppose it happens to be the case that the day before, November 26, was cold and rainy, and that the probability for the immediate successors (S) of cold and rainy days in Stanford to be warm and sunny is .2; then the account (3m) has a rival in the following argument which, by reference to equally true premises, presents it as fairly certain that November 27 is not warm and sunny:

$$\text{(3n)} \qquad \begin{array}{l} p\,(\overline{W}, S) = .8 \\ Sn \\ \hline\hline \overline{W}n \end{array} \quad [.8]$$

In this form, the problem of ambiguity concerns I-S arguments whose premises are in fact true, no matter whether we are aware of this or not. But, as will now be shown, the problem has a variant that concerns explanations whose explanans statements, no matter whether in fact true or not, are *asserted or accepted* by empirical science at the time when the explanation is proffered or contemplated. This variant will be called *the problem of the epistemic ambiguity of statistical explanation*, since it refers to what is presumed to be known in science rather than to what, perhaps unknown to anyone, is in fact the case.

Let K_t be the class of all statements asserted or accepted by empirical science at time t. This class then represents the total scientific information, or "scientific knowledge" at time t. The word 'knowledge' is here used in the sense in which we commonly speak of the scientific knowledge at a given time. It is not meant to convey the claim that the elements of K_t are true, and hence neither that they are definitely known to be true. No such claim can justifiably be made for any of the statements established by empirical science; and the basic standards of scientific inquiry demand that an empirical statement, however well supported, be accepted and thus admitted to membership in K_t only tentatively, i.e., with the understanding that the privilege may be withdrawn if unfavorable evidence should be discovered. The membership of K_t therefore changes in the course of time; for as a result of continuing research, new statements are admitted into that class; others may come to be discredited and dropped. Henceforth, the class of accepted statements will be referred to simply as K when specific reference to the time in question is not required. We will assume that K is logically consistent and that it is closed under logical implication, i.e., that it contains every statement that is logcially implied by any of its subsets.

The *epistemic ambiguity of I-S explanation* can now be characterized as follows: The total set K of accepted scientific statements contains different subsets of statements which can be used as premises in arguments of the probabilistic form just considered, and which confer high probabilities on logically contradictory "conclusions." Our earlier examples (3k), (3l) and (3m), (3n) illustrate this point if we assume that the premises of those arguments all belong to K rather than that they are all true. If one of two such rival arguments with premises in K is proposed as an explanation of an event considered, or acknowledged, in science to have occurred, then the conclusion of the argument, i.e., the explanandum statement, will accordingly belong to K as well. And since K is consistent, the conclusion of the rival argument will not belong to K. Nonetheless it is disquieting that we should be able to say: No matter whether we are informed that the event in question (e.g., warm and sunny weather on November 27 in Stanford) did occur or that it did not occur, we can produce an explanation of the reported outcome in either case; and an explanation, moreover, whose premises are scientifically established statements that confer a high logical probability upon the reported outcome.

This epistemic ambiguity, again, has no analogue for deductive explanation; for since K is logically consistent, it cannot contain premise-sets that imply logically contradictory conclusions.

Epistemic ambiguity also bedevils the predictive use of statistical arguments. Here, it has the alarming aspect of presenting us with two rival arguments whose premises are scientifically well established, but one of which characterizes a contemplated future occurrence as practically certain, whereas the other characterizes it as practically impossible. Which of such conflicting arguments, if any, are rationally to be relied on for explanation or for prediction?

3.4.2 *The Requirement of Maximal Specificity and the Epistemic Relativity of Inductive-Statistical Explanation.* Our illustrations of explanatory ambiguity suggest that a decision on the acceptability of a proposed probabilistic explanation or prediction will have to be made in the light of all the relevant information at our disposal. This is indicated also by a general principle whose importance for inductive reasoning has been acknowledged, if not always very explicitly, by many writers, and which has recently been strongly emphasized by Carnap, who calls it *the requirement of total evidence*. Carnap formulates it as follows: "in the application of inductive logic to a given knowledge situation, the total evidence available must be taken as basis for determining the degree of confirmation."[17] Using only a part of the total evidence is permissible if the balance of the evidence is irrelevant to the inductive "conclusion," i.e., if on the partial evidence alone, the conclusion has the same confirmation, or logical probability, as on the total evidence.[18]

The requirement of total evidence is not a postulate nor a theorem of inductive logic; it is not concerned with the formal validity of inductive arguments. Rather, as Carnap has stressed, it is a maxim for the *application* of inductive logic; we might say that it states a necessary condition of rationality of any such application in a given "knowledge situation," which we will think of as represented by the set K of all statements accepted in the situation.

But in what manner should the basic idea of this requirement be brought to bear upon probabilistic explanation? Surely we should not insist that the explanans must contain all and only the empirical information available at the time. Not *all* the available information, because otherwise all probabilistic explanations acceptable at time t would have to have the same explanans, K_t; and not *only* the available information, because a proffered explanation may

17. Carnap (1950), p. 211.

The requirement is suggested, for example, in the passage from Lewis (1946) quoted in note 5 for this section. Similarly Williams speaks of "the most fundamental of all rules of probability logic, that 'the' probability of any proposition is its probability in relation to the known premises and them only." (Williams, 1947, p. 72).

I am greatly indebted to Professor Carnap for having pointed out to me in 1945, when I first noticed the ambiguity of probabilistic arguments, that this was but one of several apparent paradoxes of inductive logic that result from disregard of the requirement of total evidence.

Barker (1957), pp. 70-78, has given a lucid independent presentation of the basic ambiguity of probabilistic arguments, and a skeptical appraisal of the requirement of total evidence as a means of dealing with the problem. However, I will presently suggest a way of remedying the ambiguity of probabilistic explanation with the help of a rather severely modified version of the requirement of total evidence. It will be called the requirement of maximal specificity, and is not open to the same criticism.

18. *Cf.* Carnap (1950), p. 211 and p. 494.

meet the intent of the requirement in not overlooking any relevant information available, and may nevertheless invoke some explanans statements which have not as yet been sufficiently tested to be included in K_t.

The extent to which the requirement of total evidence should be imposed upon statistical explanations is suggested by considerations such as the following. A proffered explanation of Jones's recovery based on the information that Jones had a streptococcal infection and was treated with penicillin, and that the statistical probability for recovery in such cases is very high is unacceptable if K includes the further information that Jones's streptococci were resistant to penicillin, or that Jones was an octogenarian with a weak heart, and that in these reference classes the probability of recovery is small. Indeed, one would want an acceptable explanation to be based on a statistical probability statement pertaining to the narrowest reference class of which, according to our total information, the particular occurrence under consideration is a member. Thus, if K tells us not only that Jones had a streptococcus infection and was treated with penicillin, but also that he was an octogenarian with a weak heart (and if K provides no information more specific than that) then we would require that an acceptable explanation of Jones's response to the treatment be based on a statistical law stating the probability of that response in the narrowest reference class to which our total information assigns Jones's illness, i.e., the class of streptococcal infections suffered by octogenarians with weak hearts.[19]

Let me amplify this suggestion by reference to our earlier example concerning the use of the law that the half-life of radon is 3.82 days in accounting for the fact that the residual amount of radon to which a sample of 10 milligrams was reduced in 7.64 days was within the range from 2.4 to 2.6 milligrams. According to present scientific knowledge, the rate of decay of a radioactive element depends solely upon its atomic structure as characterized by its atomic number and its mass number, and it is thus unaffected by the age of the sample and by such factors as temperature, pressure, magnetic and electric forces, and chemical interactions. Thus, by specifying the half-life of radon as well as the initial mass of the sample and the time interval in question, the explanans takes into account all the available information that is relevant to

19. This idea is closely related to one used by Reichenbach (*cf.* (1949), section 72) in an attempt to show that it is possible to assign probabilities to individual events within the framework of a strictly statistical conception of probability. Reichenbach proposed that the probability of a single event, such as the safe completion of a particular scheduled flight of a given commercial plane, be construed as the statistical probability which the *kind* of event considered (safe completion of a flight) possesses within the narrowest reference class to which the given case (the specified flight of the given plane) belongs, and for which reliable statistical information is available (for example, the class of scheduled flights undertaken so far by planes of the line to which the given plane belongs, and under weather conditions similar to those prevailing at the time of the flight in question).

appraising the probability of the given outcome by means of statistical laws. To state the point somewhat differently: Under the circumstances here assumed, our total information K assigns the case under study first of all to the reference class, say F_1, of cases where a 10 milligram sample of radon is allowed to decay for 7.68 days; and the half-life law for radon assigns a very high probability, within F_1, to the "outcome," say G, consisting in the fact that the residual mass of radon lies between 2.4 and 2.6 milligrams. Suppose now that K also contains information about the temperature of the given sample, the pressure and relative humidity under which it is kept, the surrounding electric and magnetic conditions, and so forth, so that K assigns the given case to a reference class much narrower than F_1, let us say, $F_1F_2F_3 \ldots F_n$. Now the theory of radioactive decay, which is equally included in K, tells us that the statistical probability of G within this narrower class is the same as within G. For this reason, it suffices in our explanation to rely on the probability $p(G,F_1)$.

Let us note, however, that "knowledge situations" are conceivable in which the same argument would not be an acceptable explanation. Suppose, for example, that in the case of the radon sample under study, the amount remaining one hour before the end of the 7.68 day period happens to have been measured and found to be 2.7 milligrams, and thus markedly in excess of 2.6 milligrams—an occurrence which, considering the decay law for radon, is highly improbable, but not impossible. That finding, which then forms part of the total evidence K, assigns the particular case at hand to a reference class, say F^*, within which, according to the decay law for radon, the outcome G is highly improbable since it would require a quite unusual spurt in the decay of the given sample to reduce the 2.7 milligrams, within the one final hour of the test, to an amount falling between 2.4 and 2.6 milligrams. Hence, the additional information here considered may not be disregarded, and an explanation of the observed outcome will be acceptable only if it takes account of the probability of G in the narrower reference class, i.e., $p(G,F_1F^*)$. (The theory of radioactive decay implies that this probability equals $p(G,F^*)$, so that as a consequence the membership of the given case in F_1 need not be explicitly taken into account.)

The requirement suggested by the preceding considerations can now be stated more explicitly; we will call it the *requirement of maximal specificity for inductive-statistical explanations*. Consider a proposed explanation of the basic statistical form

(30)
$$\begin{array}{l} p(G,F) = r \\ Fb \\ \hline\hline Gb \end{array} \; [r]$$

Let s be the conjunction of the premises, and, if K is the set of all statements accepted at the given time, let k be a sentence that is logically equivalent to K (in the sense that k is implied by K and in turn implies every sentence in K). Then, to be rationally acceptable in the knowledge situation represented by K, the proposed explanation (30) must meet the following condition (the requirement of maximal specificity): If $s \cdot k$ implies[20] that b belongs to a class F_1, and that F_1 is a subclass of F, then $s \cdot k$ must also imply a statement specifying the statistical probability of G in F_1, say

$$p(G, F_1) = r_1$$

Here, r_1 must equal r unless the probability statement just cited is simply a theorem of mathematical probability theory.

The qualifying unless-clause here appended is quite proper, and its omission would result in undesirable consequences. It is proper because theorems of pure mathematical probability theory cannot provide an explanation of empirical subject matter. They may therefore be discounted when we inquire whether $s \cdot k$ might not give us statistical laws specifying the probability of G in reference classes narrower than F. And the omission of the clause would prove troublesome, for if (30) is proffered as an explanation, then it is presumably accepted as a fact that Gb; hence 'Gb' belongs to K. Thus K assigns b to the narrower class $F \cdot G$, and concerning the probability of G in that class, $s \cdot k$ trivially implies the statement that $p(G, F \cdot G) = 1$, which is simply a consequence of the measure-theoretical postulates for statistical probability. Since $s \cdot k$ thus implies a more specific probability statement for G than that invoked in (30), the requirement of maximal specificity would be violated by (30)—and analogously by any proffered statistical explanation of an event that we take to have occurred—were it not for the unless-clause, which, in effect, disqualifies the notion that the statement '$p(G, F \cdot G) = 1$' affords a more appropriate law to account for the presumed fact that Gb.

The requirement of maximal specificity, then, is here tentatively put forward as characterizing the extent to which the requirement of total evidence properly applies to inductive-statistical explanations. The general idea thus suggested comes to this: In formulating or appraising an I-S explanation, we should take into account all that information provided by K which is of potential *explanatory* relevance to the explanandum event; i.e., all pertinent statistical laws, and such

20. Reference to $s \cdot k$ rather than to k is called for because, as was noted earlier, we do not construe the condition here under discussion as requiring that all the explanans statements invoked be scientifically accepted at the time in question, and thus be included in the corresponding class K.

particular facts as might be connected, by the statistical laws, with the explanandum event.[21]

The requirement of maximal specificity disposes of the problem of epistemic ambiguity; for it is readily seen that of two rival statistical arguments with high associated probabilities and with premises that all belong to K, at least one violates the requirement of maximum specificity. Indeed, let

$$\begin{array}{c} p\,(G,F) = r_1 \\ F\,b \\ \hline\hline G\,b \end{array}\ [r_1] \qquad \text{and} \qquad \begin{array}{c} p(\overline{G}, H) = r_2 \\ H\,b \\ \hline\hline \overline{G}\,b \end{array}\ [r_2]$$

be the arguments in question, with r_1 and r_2 close to 1. Then, since K contains the premises of both arguments, it assigns b to both F and H and hence to $F{\cdot}H$. Hence if both arguments satisfy the requirement of maximal specificity, K must imply that

$$p(G, F{\cdot}H) = p(G, F) = r_1$$
$$p(\overline{G}, F{\cdot}H) = p(\overline{G}, H) = r_2$$

But
$$p(G, F{\cdot}H) + p(\overline{G}, F{\cdot}H) = 1$$

Hence
$$r_1 + r_2 = 1$$

and this is an arithmetic falsehood, since r_1 and r_2 are both close to 1; hence it cannot be implied by the consistent class K.

Thus, for I-S explanations that meet the requirement of maximal specificity the problem of epistemic ambiguity no longer arises. We are *never* in a position to say: No matter whether this particular event did or did not occur, we can produce an acceptable explanation of either outcome; and an explanation, moreover, whose premises are scientifically accepted statements which confer a high logical probability upon the given outcome.

21. By its reliance on this general idea, and specifically on the requirement of maximal specificity, the method here suggested for eliminating the epistemic ambiguity of statistical explanation differs substantially from the way in which I attempted in an earlier study (Hempel, 1962, especially section 10) to deal with the same problem. In that study, which did not distinguish explicitly between the two types of explanatory ambiguity characterized earlier in this section, I applied the requirement of total evidence to statistical explanations in a manner which presupposed that the explanans of any acceptable explanation belongs to the class K, and which then demanded that the probability which the explanans confers upon the explanandum be equal to that which the total evidence, K, imparts to the explanandum. The reasons why this approach seems unsatisfactory to me are suggested by the arguments set forth in the present section. Note in particular that, if strictly enforced, the requirement of total evidence would preclude the possibility of any significant statistical explanation for events whose occurrence is regarded as an established fact in science; for any sentence describing such an occurrence is logically implied by K and thus trivially has the logical probability 1 relative to K.

While the problem of epistemic ambiguity has thus been resolved, ambiguity in the first sense discussed in this section remains unaffected by our requirement; i.e., it remains the case that for a given statistical argument with true premises and a high associated probability, there may exist a rival one with equally true premises and with a high associated probability, whose conclusion contradicts that of the first argument. And though the set K of statements accepted at any time never includes all statements that are in fact true (and no doubt many that are false), it is perfectly possible that K should contain the premises of two such conflicting arguments; but as we have seen, at least one of the latter will fail to be rationally acceptable because it violates the requirement of maximal specificity.

The preceding considerations show that *the concept of statistical explanation for particular events is essentially relative to a given knowledge situation as represented by a class K of accepted statements*. Indeed, the requirement of maximal specificity makes explicit and unavoidable reference to such a class, and it thus serves to characterize the concept of "I-S explanation relative to the knowledge situation represented by K." We will refer to this characteristic as the *epistemic relativity of statistical explanation*.

It might seem that the concept of deductive explanation possesses the same kind of relativity, since whether a proposed D-N or D-S account is acceptable will depend not only on whether it is deductively valid and makes essential use of the proper type of general law, but also on whether its premises are well supported by the relevant evidence at hand. Quite so; and this condition of empirical confirmation applies equally to statistical explanations that are to be acceptable in a given knowledge situation. But the epistemic relativity that the requirement of maximal specificity implies for I-S explanations is of quite a different kind and has no analogue for D-N explanations. For the specificity requirement is not concerned with the evidential support that the total evidence K affords for the explanans statements: it does not demand that the latter be included in K, nor even that K supply supporting evidence for them. It rather concerns what may be called the concept of a *potential* statistical explanation. For it stipulates that no matter how much evidential support there may be for the explanans, a proposed I-S explanation is not acceptable if its potential explanatory force with respect to the specified explanandum is vitiated by statistical laws which are included in K but not in the explanans, and which might permit the production of rival statistical arguments. As we have seen, this danger never arises for deductive explanations. Hence, these are not subject to any such restrictive condition, and the notion of a potential deductive explanation (as contradistinguished from a deductive explanation with well-confirmed explanans) requires no relativization with respect to K.

As a consequence, we can significantly speak of true D-N and D-S explanations: they are those potential D-N and D-S explanations whose premises (and hence also conclusions) are true—no matter whether this happens to be known or believed, and thus no matter whether the premises are included in *K*. But this idea has no significant analogue for I-S explanation since, as we have seen, the concept of potential statistical explanation requires relativization with respect to *K*.

3.4.3 *Discrete State Systems and Explanatory Ambiguity*. In a lucid and instructive article, Rescher[22] has shown that physical systems of a particular kind, which he calls discrete state systems, afford excellent illustrations of deductive and probabilistic explanation and prediction, and that a closer examination of such systems can shed a good deal of light on the logical structure, the scope, and the interrelations of those procedures. I propose to show that a study of those systems also confronts one with the problem of explanatory ambiguity and supports the solution here suggested.

By a *discrete state system*, or a DS system for short, Rescher understands a physical system which at any moment is in one of several possible states, $S_1, S_2, \ldots$, each of whose occurrences occupies a finite, though perhaps very brief, period of time; for the purpose at hand, the number of possible states for a DS system is taken to be finite. The succession of states exhibited by a DS system is governed by a set of laws, each of which may be deterministic or probabilistic (statistical). A deterministic law has the form 'State S_i is always immediately followed by state S_j'; a probabilistic law has the form 'The statistical probability for (an occurrence of) state S_i to be immediately followed by (an occurrence of) state S_j is r_{ij}.' A DS system of this kind can be characterized by means of the matrix of all the transition probabilities r_{ij}.

There are various physical examples of DS systems; among them Rescher mentions an electronic digital computer; an atom of a radioactive element in its successive states of decay; and—given a suitably schematized mode of description—a particle in Brownian motion. A ball rolling down a Galton Board[23] is yet another DS system; its state at a given time being represented by the number of pins that separate it horizontally from the vertical center line of the board.

A potential probabilistic explanation (of a momentary state of a DS system) is defined by Rescher as an argument whose conclusion is of the form 'the state of the system in time-interval t is S_i', or '$st(t) = S_i$' for short, and whose premises consist of the laws governing the system and of a set of statements specifying the states exhibited by the system during certain other time intervals,

22. Rescher (1963).

23. For a discussion of this process, see Mises (1939), pp. 237–40.

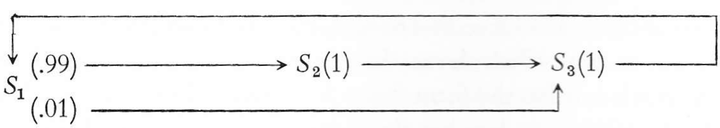
$t_1, t_2, \ldots, t_n$, all of which are different from t.[24] The argument may be "probabilistic, either in the strong sense. . . that $st(t) = S_i$ is (conditionally) *more likely than not*, or in the weak sense . . . that $st(t) = S_i$ is (conditionally) *more likely than* $st(t) = S_j$ for any $j \neq i$."[25] Finally, "A potentially explanatory argument becomes an (*actual*) *explanation* if its premises are actually or probably true."[26]

To see that probabilistic explanation thus construed again is plagued by ambiguity, consider a DS system capable of just three states, S_1, S_2, S_3, with transition probabilities as specified in the following schema:

	S_1	S_2	S_3
S_1	0	.99	.01
S_2	0	0	1
S_3	1	0	0

Thus, the probability of S_1 being immediately followed by S_1 is 0; by S_2, .99; by S_3, .01; and so forth.

Alternatively, DS systems can be characterized by what Rescher calls transition-diagrams. In our case, the diagram takes the following form:



As is readily seen, the transition laws here indicated imply the following two derivative laws:

(L_1) The probability for the two-period successor of S_1 to be S_3 is $.99 \times 1 = .99$.

(L_2) The probability for the immediate successor of S_3 to be again S_3 is 0.

Suppose now that in two particular successive time intervals t_1 and t_2, our

24. Rescher does not require of a potential explanation—as he does of a potential prediction, which is otherwise characterized in the same manner—that the time intervals t_1, t_2, . . . , t_n must all precede t. As a result, every potential prediction is a potential explanation, but not conversely. His reason for this construal will be examined in section 3.5.

25. Rescher (1963), p. 330, italics the author's. The concept of conditional likelihood here invoked is not further clarified; but it evidently is meant to represent the likelihood which the conclusion of the explanatory argument possesses relative to, or conditional upon, the premises. In this case, likelihoods would have the general character of logical probabilities; and Rescher does seem to operate with them in accordance with the conception reflected by our schema (3h), where the "likelihood" in question is specified in square brackets next to the double line separating the conclusion from the premises.

26. Rescher (1963), p. 329, italics the author's.

system exhibits the states S_1 and S_3 respectively; i.e. that the following statements are true:

(C_1) $st(t_1) = S_1$
(C_2) $st(t_2) = S_3$

Then C_1 jointly with L_1 provides the premises for a probabilistic argument which gives the "likelihood" .99 to the conclusion that in the time interval t_3 immediately following t_2, the system is in state S_3; i.e., that $st(t_3) = S_3$. But C_2 jointly with L_2 analogously gives the likelihood 1 to the conclusion that $st(t_3) \neq S_3$. On our assumptions, the premises invoked in these conflicting arguments are true; hence the arguments constitute strong probabilistic explanations, in Rescher's sense, of the occurrence and of the nonoccurrence of S_3 during time interval t_3; and both are actual explanations in Rescher's sense since all the explanatory premises are true. Thus we have explanatory ambiguity in the first of our two senses. That ambiguity in the second, epistemic, sense is present as well is clear when we consider that on our assumptions, all the premises invoked may of course belong to the class K of statements that are accepted at the time.[27]

To preclude this untenable consequence, Rescher's definitions of probabilistic explanation and prediction must be supplemented by a suitable additional requirement. In our example, the first of the two competing arguments would clearly be rejected on the ground that it disregards some relevant information. But this is precisely the verdict of the requirement of maximal specificity. For in our illustration, we may assume that the class K includes the information conveyed by C_1, C_2, L_1, and L_2; but that K contains no more specific information which would imply a probability assignment, on empirical grounds, to the sentence '$st(t_3) = S_3$'. The first of the two probabilistic arguments violates the requirement of maximal specificity, since it takes into account only that the state of the system at t_1 is S_1, although K tells us further that the occurrence of S_1 at t_1 is directly followed by an occurrence of S_2, and that for an occurrence of S_1 that is followed by an occurrence of S_2 the probability of having S_3 as a two-period successor is 0. (For L_2 tells us quite generally that the probability for an occurrence of S_2—no matter what its predecessor may be—to be followed by an occurrence of S_3 is 0.) Hence only the second of the two rival arguments is acceptable under the requirement of maximal specificity.

27. The same ambiguity would jeopardize the predictive use of these arguments: though both based on accepted (and indeed, true) premises, they lead to contradictory predictions about the state of the system during t_3.

3.5 PREDICTIVE ASPECTS OF STATISTICAL EXPLANATION. Can it be maintained that an inductive-statistical explanation of a particular event, much like a deductive-nomological one, constitutes a potential prediction of that event?

If the statement describing the occurrence in question is included in the class K of accepted statements, then the question of predicting the event clearly cannot arise in the knowledge situation represented by K. Let us therefore put our problem into this form: Suppose that an argument of the type (30) meets the requirement of maximum specificity relative to K and that its explanans is well confirmed by K; would it then be acceptable as a predictive argument in the knowledge situation characterized by K? The answer will depend, of course, on the conditions we think a statistical argument has to satisfy if it is to be rationally acceptable for predictive purposes in a given knowledge situation. Let us briefly consider this question.

Rationality clearly demands that in forming expectations concerning future occurrences we take into account all the relevant information available at the time: this is the gist of the requirement of total evidence. But how is this requirement to be construed more specifically? If a general definition and theory of logical, or inductive, probability is available, the condition comes to this: the probability conferred upon the conclusion of the predictive argument by the premises alone should equal the probability imparted to it by the total evidence K; in that case, the balance of the total evidence is justifiably disregarded in the argument, for its addition to the premises would not change the probability of the conclusion. At present, no definition and theory of inductive probability is available which is sufficiently comprehensive to be applicable to all the kinds of inductive argument that would have to be considered. If such a definition should be constructed—for example, by generalizing Carnap's approach—it might turn out that a statistical argument whose premises are well supported by K, and which does satisfy the requirement of maximal specificity, still does not strictly meet the requirement of total evidence in the precise quantitative form under consideration. For example, let K consist of the premises of (30) and the further statement 'Hd', then, though intuitively this latter statement is entirely irrelevant to the conclusion 'Gb', it is conceivable that the logical probability, in the sense here assumed, of 'Gb' relative to K should differ from the logical probability r of 'Gb' relative to the premises of (30) alone. Or suppose that K consists of the statements '$p(G,F) = .9$', '$p(G,H)=.1$', '$p(G,F{\cdot}H)=.85$', 'Fb', 'Hb'; then a statistical argument with the last three of these statements as premises and 'Gb' as conclusion satisfies the requirement of maximal specificity relative to K. Yet again, the logical probability of 'Gb' relative to K might differ from the logical probability, .85, of 'Gb' relative to the set of the three premise-statements.

In the absence of a suitable general definition of logical probability, however, it seems quite clear that the predictive argument just considered would indeed by regarded as rationally acceptable in the knowledge situation represented by K; the statistical law specifying the probability of G in $F \cdot H$ would count as overriding the laws specifying the probability of G relative to F and to H, respectively. Similarly, an argument of the type (30) whose premises are well substantiated and which conforms to the requirement of maximum specificity would surely be regarded as a rational way of forming expectations concerning the event described by the conclusion. And in general, predictive arguments in science which are based on probabilistic laws appear to be governed by the requirement of maximum specificity and the requirement of adequate confirmation for the premises. To the extent thus indicated, then, an argument that constitutes an acceptable statistical explanation relative to K also forms an acceptable potential prediction relative to K.

Hanson[28] has put forward an interesting view of the relation between explanatory and predictive arguments in science, which gives me an occasion as well as an opportunity to amplify the general position just outlined, and to argue further in its support.

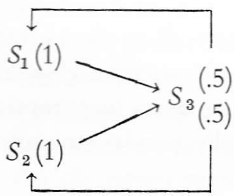
According to Hanson, the view that an adequate explanation also affords a potential prediction conforms well to the character of the explanations and predictions made possible by Newtonian classical mechanics, which is deterministic in character; but it is quite inappropriate in reference to quantum theory, which is fundamentally nondeterministic. More specifically, Hanson holds that the laws of quantum theory do not permit the *prediction* of any individual quantum phenomenon P, such as the emission of a beta-particle by a radioactive substance, but that "P can be completely *explained* ex post facto; one can understand fully just what kind of event occurred, in terms of the well-established laws of . . . quantum theory. . . . These laws give the *meaning* of 'explaining single microevents'."[29]

It is indeed the case that because of their purely statistical character, the laws of radioactive decay permit the prediction of events such as the emission of beta-particles by disintegrating atoms only with probability and not with deductive-nomological definiteness for an individual occurrence. But for exactly the same reason, those laws permit only a probabilistic explanation of a particular emission P rather than a "complete" explanation "ex post facto," as Hanson puts it. For if, as the phrase "ex post facto" might seem to suggest, the information that P has occurred were included in the explanans, the

28. Hanson (1959) and (1963), chapter 2.
29. Hanson (1959), p. 354, italics the author's; similarly in Hanson (1963), p. 29.

resulting account would be unilluminatingly circular: surely Hanson does not mean that. And if the explanans contains only statements about antecedent conditions, plus the statistical laws of radioactive decay, then it can show at best that the occurrence of P was highly probable; but this affords only an inductive-statistical explanation, which has the same logical form as the probabilistic, i.e., inductive-statistical, prediction of P.[30]

In the context of his argument, Hanson puts forward another assertion, namely: "Every prediction, if inferentially respectable, must possess a corresponding postdiction."[31] By a postdiction, Hanson means "simply the logical reversal of a prediction": a prediction proceeds "from initial conditions through boundary conditions to a statement about some future event x," and a postdiction consists "in inferring from a statement about some present event x, through the boundary conditions, back to already *known* initial conditions."[32] But Hanson's thesis is incorrect, as is shown by the following counter-example. Consider a discrete state system whose three possible states, S_1, S_2, S_3, are linked by the following laws: S_1 as well as S_2 is always followed by S_3; S_3 is followed, with a probability of .5, by S_1 and with the same probability by S_2. The corresponding transition diagram is this:



Then the information that in time-interval t_5 the system is in S_2 permits the deductive-nomological, and thus clearly "inferentially respectable" prediction that during t_6, the system will be in S_3; but no corresponding postdiction is possible from the latter information to the former.[33]

In conclusion, I wish to consider an argument put forward by Rescher as to the relation between explanation and prediction. The gist of it can most

30. For comments in a similar vein, see Henson (1963); *cf.* also the critical response in Feyerabend (1964).

31. Hanson (1963), p. 193, *cf.* also p. 40. Hanson goes on to say: "This is part of Hempel's thesis, and it is sound, necessarily" (*Ibid.*). Actually, I have argued *against* this thesis, which is true of predictions based on deterministic theories, but not true in general. See Hempel (1962), pp. 114–15.

32. Hanson (1963), p. 193, italics the author's.

33. On this point, see also Grünbaum (1963), p. 76. Grünbaum's article presents a detailed discussion of the structural identity of explanation and prediction and examines a variety of objections to this idea.

simply be stated by particular reference to Rescher's study of discrete state systems, which we considered in section 3.4. On Rescher's definition, an argument explaining the state of such a system in time interval t may refer, in the explanans, to the states exhibited by the system at certain other times, which may be earlier or later than t; whereas an argument predicting the state at t is required to refer only to preceding states. As a consequence of these definitional stipulations, "it follows that whenever a prediction . . . is given, so *a fortiori* is an explanation," but not conversely. "For our defining conditions for prediction . . . in effect add to the conditions for explanation . . . certain added restrictions of a temporal character."[34]

In defense of imposing that additional requirement on prediction, Rescher argues, in effect, as follows: Suppose that the premises of a proposed argument predicting the state of the system at t include a statement specifying the state of that system for some later time interval t_1. Then, since the argument is predictive, t is later than "the present," t_N, and hence so is t_1. Now there are two possibilities. Either (i) the premise pertaining to t_1 can itself be inferred, by means of laws, from past states of the system: then the given predictive argument can evidently be replaced by one that infers the state at t, with the help of laws, solely from past states, so that the restrictive requirement is met; or (ii) the explanatory premise about t_1 cannot be inferred from statements about past states: then "we do not actually have a proper prediction at all —for we are basing our 'predictive' argument on a premise which cannot be justified in terms of *available* information."[35]

But as the reference to justificatory evidence indicates, this consideration has no bearing on the thesis that an explanatory argument is potentially also a predictive one, i.e., that it could have been used to derive a predictive sentence concerning the state of the system at t if the statements forming the explanans had been formulated and used as premises before t. To be sure, we would normally ask for an explanation of a given state only after its occurrence, i.e., in our case, after t;[36] and it is true, as the argument points out, that we may then be able to support the critical premise by evidence that was not available before t. But the empirical support for the premises has no bearing on the structural relationships between explanatory and predictive arguments; nor, I think, do considerations based on it afford good grounds for imposing a restrictive formal condition upon predictive inferences.

34. Rescher (1963), p. 329.

35. Rescher (1963), p. 333, italics the author's.

36. Indeed, by parity of reasoning, Rescher would seem obliged to say that the argument considered in our example, one of whose premises refers to t_1, is not a proper explanation either, if it is presented before t_1 (though after t), for it then rests on a premise that is not justified by available evidence.

It should also be remembered that, as was noted in section 2.4, even the most perfect cases of scientific prediction normally make use of some statements about the future that are not inferred by law from information about the past. Thus, the prediction of the positions of the planets at a given time on the basis of the requisite data concerning their locations and momenta a month earlier requires an assumption concerning the boundary conditions during the intervening time interval, normally to the effect that there will be no outside interference with the system. And though this is not inferred by law from other particulars, the arguments presupposing those boundary conditions are not regarded as therefore affording no proper predictions at all.

Finally, we might note with Scheffler that we may sometimes reasonably speak of explaining a future event, and that indeed, in some cases, one and the same argument may be considered as predicting a certain event and explaining it; as, for example, when the question 'Why will the sun rise tomorrow?' is answered by offering some appropriate astronomical information.[37] For this reason, too, it seems inadvisable to impose different formal requirements upon explanatory and predictive arguments.

3.6 THE NONCONJUNCTIVENESS OF INDUCTIVE-STATISTICAL EXPLANATION. Inductive-statistical explanation differs from its deductive counterparts in yet another important respect. When a given explanans deductively accounts for each of several explananda, then it also deductively accounts for their conjunction; but the analogue for I-S explanation does not generally hold because an explanans that confers high probability on each of several explananda may confer a very low probability on their conjunction. In this sense, then, *I-S explanation, in contrast to deductive explanation, is non-conjunctive.*

Consider, for example, the random experiment F of flipping a fair coin ten times in succession. Each performance of this experiment will yield, as its outcome, one of the $2^{10} = 1024$ different possible sequences of 10 individual results each of which is either heads or tails. Let $O_1, O_2, \ldots, O_{1024}$ be the different possible kinds of outcome thus characterized. Then, according to the standard statistical hypothesis—let us call it S—for this kind of experiment, the probability of obtaining heads by flipping the coin is 1/2, and the results of different flippings are statistically independent of each other. It follows therefore deductively that the statistical probability of obtaining outcome O_k as a result of performing F is $p(O_k, F) = 1/1024$, and the probability of getting a result other than O_k is $p(\overline{O}_k, F) = 1\text{-}1/1024 = 1023/1024$, for any one of the different possible outcomes O_k.

37. Scheffler (1957), p. 300.

Suppose now that a particular performance, f, of F has yielded O_{500} as its outcome: $O_{500}(f)$. This result can also be described by saying that f did not yield any of the other possible outcomes:

$$\bar{O}_1(f) \cdot \bar{O}_2(f) \ldots \bar{O}_{499}(f) \cdot \bar{O}_{501}(f) \cdot \ldots \cdot \bar{O}_{1024}(f)$$

Now, our statistical hypothesis S in conjunction with the information that f was a particular performance of F, i.e., that $F(f)$, provides an I-S explanation with high associated probability for (the facts described by) each of the 1023 sentences here conjoined as follows:

$$\begin{array}{ll} p(\bar{O}_k, F) = 1023/1024 & \\ F(f) & \\ \hline\hline & [1023/1024] \\ \bar{O}_k(f) & \end{array}$$

The requirement of maximal specificity is satisfied by these accounts since for such further information about the particular experiment f as may be available to us under the circumstances, S may be taken to imply that it does not affect the probability of O_k. But though S in combination with the information that $F(f)$ thus confers a high probability on each of the 1023 conjoined statements just listed, it assigns the very low probability of 1/1024 to their conjunction, which is tantamount to the statement '$O_{500}(f)$';
For we have

$$\begin{array}{ll} p(O_{500}, F) = 1/1024 & \\ F(f) & \\ \hline\hline & [1/1024] \\ O_{500}(f) & \end{array}$$

Thus, while S together with '$F(f)$' provides an I-S explanation with high associated probability for (the facts described by) any of the 1023 sentences cited above, it does not do so for (the facts described by) their conjunction.[38]

This nonconjunctiveness of I-S explanation thus springs from the fact that one and the same set of sentences may confirm to a very high degree each of n alternative statements while confirming with similar strength the negation of their conjunction. This fact, in turn, is rooted in the general multiplication theorem for probabilities, which implies that the probability of the conjunction of two items (i.e., characteristics or sentences, according as statistical or logical probabilities are concerned) is, in general, less than the probability of either item taken by itself. Hence, once the connection between explanans and explanandum in the statistical explanation of particular phenomena is viewed as

38. For another illustration, *cf.* Hempel (1962), p. 165.

inductive, nonconjunctiveness presents itself as an inevitable aspect of it, and thus as one of the fundamental characteristics that set I-S explanation apart from its deductive counterparts.

4. THE CONCEPTS OF COVERING-LAW EXPLANATION AS EXPLICATORY MODELS

4.1 General Character and Intent of the Models. We have by now distinguished three basic types of scientific explanative: deductive-nomological, inductive-statistical, and deductive-statistical. The first of these is often referred to as the covering-law model or the deductive model of explanation, but since the other two types also involve reference to covering laws, and since one of them is deductive as well, we will call the first more specifically the *deductive-nomological model;* analogously, we will speak of the others as the *inductive-statistical* and the *deductive statistical models of explanation.*

As is made clear by our earlier discussions, these models are not meant to describe how working scientists actually formulate their explanatory accounts. Their purpose is rather to indicate in reasonably precise terms the logical structure and the rationale of various ways in which empirical science answers explanation-seeking why-questions. The construction of our models therefore involves some measure of abstraction and of logical schematization.

In these respects, our concepts of explanation resemble the concept, or concepts, of mathematical proof (within a given mathematical theory) as construed in metamathematics. Let us note the principal points of resemblance.

In either case, the models seek to explicate the use and function of certain "explicandum" terms—'proof' and its cognates in one case, 'explanation' and its cognates in the other. However, the models are selective; they are not meant to illuminate all the different customary uses of the terms in question, but only certain special ones. Thus, metamathematical proof theory is concerned only with the notion of proof in mathematics. To put the theory forward is not to deny that there are other contexts in which we speak of proofs and proving, nor is it to assert that the metamathematical concepts are relevant to those contexts.

Similarly, to put forward the covering-law models of scientific explanation is not to deny that there are other contexts in which we speak of explanation, nor is it to assert that the corresponding uses of the word 'explain' conform to one or another of our models. Obviously, those models are not intended to reflect the various senses of 'explain' that are involved when we speak of explaining the rules of a contest, explaining the meaning of a cuneiform inscription or of a complex legal clause or of a passage in a symbolist poem,