Introduction

The search for scientific knowledge extends far back into antiquity. At some point in that quest, at least by the time of Aristotle, philosophers recognized that a fundamental distinction should be drawn between two kinds of scientific knowledge—roughly, knowledge *that* and knowledge *why*. It is one thing to know *that* each planet periodically reverses the direction of its motion with respect to the background of fixed stars; it is quite a different matter to know *why*. Knowledge of the former type is descriptive; knowledge of the latter type is explanatory. It is explanatory knowledge that provides scientific understanding of our world.

Nevertheless, when Aristotle and many of his successors down through the centuries tried to say with some precision what constitutes scientific explanation they did not meet with great success. According to Aristotle, scientific explanations are deductive arguments; as we shall see, this idea has been extraordinarily influential. But as Aristotle clearly recognized, not all deductive arguments can qualify as explanations. Even if one accepts the idea that explanations are deductive arguments, it is no easy matter to draw a viable distinction between those arguments that do qualify and those that do not.

Forty years ago a remarkable event occurred. Carl G. Hempel and Paul Oppenheim published an essay, "Studies in the Logic of Explanation," which was truly epoch-making. It set out, with unprecedented precision and clarity, a characterization of one kind of deductive argument that, according to their account, does constitute a legitimate type of scientific explanation. It came later to be known as *the deductive-nomological model*. This 1948 article provided the foundation for the *old consensus* on the nature of scientific explanation that reached its height in the 1960s. A large preponderance of the philosophical work on scientific explanation in the succeeding four decades has occurred as a direct

I should like to express my sincere thanks to Marc Lange for expert bibliographical assistance, and my heartfelt gratitude to Paul Humphreys, Philip Kitcher, and Nicholas Rescher for extremely valuable comments on an earlier draft of this essay. My greatest debt is to Philip Kitcher for his psychological support and intellectual stimulation, without which it would never have been written.

or indirect response to this article. If we wish to assess the prospects for a *new consensus* on scientific explanation, this is where we must start. To understand the present situation we need to see how the old consensus came together and how it came apart.

0.1 A Bit of Background

I recall with amusement a personal experience that occurred in the early 1960s. J. J. C. Smart, a distinguished Australian philosopher, visited Indiana University where I was teaching at the time. Somehow we got into a conversation about the major unsolved problems in philosophy of science, and he mentioned the problem of scientific explanation. I was utterly astonished—literally, too astonished for words. At the time I considered that problem essentially solved by the deductive-nomological (D-N) account that had been promulgated by R. B. Braithwaite (1953), Carl G. Hempel (Hempel and Oppenheim 1948), Ernest Nagel (1961), and Karl Popper (1935, 1959), among many others—supplemented, perhaps, by Hempel's then recent account of statistical explanation (Hempel 1962). Although this general view had a few rather vocal critics such as N. R. Hanson (1959) and Michael Scriven (1958, 1959, 1962, 1963) it was widely accepted by scientifically minded philosophers; indeed, it qualified handily as the received view. What is now amusing about the incident is my naïveté in thinking that a major philosophical problem had actually been solved, but my attitude did reflect the then current almost complete consensus.

On one fundamental issue the consensus has remained intact. Philosophers of very diverse persuasions continue to agree that a fundamental aim of science is to provide explanations of natural phenomena. During the last forty years, few (if any) have voiced the opinion that the sole aims of science are to describe, predict, and control nature—that explanation falls into the domains of metaphysics or theology. It has not always been so. Twentieth-century scientific philosophy arose in a philosophical context dominated by post-Kantian and post-Hegelian German idealism. It was heavily infused with transcendental metaphysics and theology. The early logical positivists and logical empiricists saw it as part of their mission to overcome such influences. As philosophers of science they were eager to expunge from science any contamination by super-empirical factors arising out of these philosophies. One such item was teleology, whether in the form of an appeal to the will of a supernatural being who created and continues to direct the course of nature, or in the form of such empirically inaccessible agencies as entelechies and vital forces. In that historical context many metaphysically inclined philosophers argued that there could be no *genuine* explanation of any fact of nature that did not involve an extra-empirical appeal. They thought of explanation anthropomorphically in terms of the sort of 'human understanding' that always appeals to purposes. Many scientific philosophers (as well as philosophical

scientists) reacted to this attitude by denying that science is in any way concerned with explanation. Those who did admit that science can offer explanations were eager to make it clear that explanation is nothing more than some special kind of description—it does not demand anything beyond the sphere of empirical knowledge.[1] The classic 1948 Hempel-Oppenheim paper, which will serve as our main point of departure, clearly illustrates this approach.

In recent decades there has been quite general agreement that science can tell us not only *what*, but also *why*. It is possible—in principle and often in practice—to furnish scientific explanations of such facts as the destruction of the space-shuttle *Challenger*, the extinction of the dinosaurs, the coppery color of the moon during total eclipse, and countless other facts, both particular and general. By means of these explanations, science provides us with genuine understanding of the world.

The philosophers who were most instrumental in forging the old consensus—the *logical empiricists*—looked upon the task of philosophy as the construction of *explications* of fundamental concepts. The clearest expression of that goal was given by Rudolf Carnap (1950, 1962, chap. 1; see also Coffa 1973). The concept we are attempting to explicate—in our case, *scientific explanation*—is known as the *explicandum*. This concept, which is frequently used by scientists and by others who talk about science, is vague and, possibly, ambiguous; the job of the philosopher is to provide a clear and exact concept to replace it. The resulting concept is known as the *explicatum*. The process of explication has two stages: first, the explicandum must be clarified sufficiently for us to know what concept it is that we are trying to explicate; second, an exact explicatum must be precisely articulated. Carnap specifies four criteria according to which explications are to be judged:

(1) Similarity to the explicandum. If the explicatum does not match the explicandum to a sufficient degree, it cannot fulfill the function of the concept it is designed to replace. A perfect match cannot, however, be demanded, for the explicandum is unclear and the explicatum should be far more pellucid.

(2) Exactness. Unless the explicatum is precise it does not fulfill the purpose of explication, namely, the replacement of an imprecise concept by a precise one.

(3) Fruitfulness. The new concept should enable us to say significant things and have important insights. One of the main benefits of philosophical analysis should be to deepen our understanding of the nature of science.

(4) Simplicity. The explicatum should be as simple as requirements (1)-(3) permit. Simplicity often accompanies systematic power of concepts. At any rate, simplicity aids in ease of application and avoidance of errors in application.

As Carnap emphatically notes, requirement (1) should not be applied too stringently. The aim is to provide a concept that is useful and clear. In the case of scientific explanation, it is evident that scientists use this concept in a variety of ways, some clear and some confused. Some scientists have claimed, for example, that explanation consists in showing how some unfamiliar phenomenon can be reduced to others that are already familiar; some have equated explanation with something that produces a feeling of intellectual satisfaction. We cannot hope, nor do we want, to capture all of these usages with complete fidelity. The logical empiricists do not indulge in 'ordinary language analysis'—even the ordinary language of scientists—except, perhaps, as a prolegomenon to philosophical analysis.

As already noted, requirement (4) is subservient to its predecessors. Thus, (2) and (3) take precedence: we seek philosophically useful concepts that are formulated with precision. Our discussion of the classic 1948 Hempel-Oppenheim paper in the next section will nicely exemplify the logical empiricist notion of explication. There are, however, several points of clarification that must be made before we turn to consideration of that paper.

First, we must be quite clear that it is *scientific* explanation with which we are concerned. The term "explanation" is used in many ways that have little or nothing to do with scientific explanation (see W. Salmon 1984, 9–11). Scriven once complained that one of Hempel's models of explanation could not even accommodate the case in which one explains with gestures what is wrong with one's car to a Yugoslav garage mechanic who knows no English. Hempel answered, entirely appropriately, that this is like complaining that a precise explication of the term "proof" in mathematics does not capture the meaning of that word as it occurs in such contexts as "86 proof Scotch" and "the proof of the pudding is in the eating" (Hempel 1965, 413). Suitable clarification of the explicandum should serve to forestall objections of that sort.

To seek an explanation for some fact presupposes, of course, that the phenomenon we endeavor to explain did occur—that the putative fact is, indeed, a fact. For example, Immanuel Velikovsky (1950) attempted to 'explain' various miracles reported in the *Old Testament*, such as the sun standing still (i.e., the earth ceasing to rotate) at Joshua's command. Those who are not dogmatically committed to the literal truth of some holy writ will surely require much stronger evidence that the alleged occurrence actually took place before surrendering such basic physical laws as conservation of angular momentum in an attempt to 'explain' it.[2]

To avoid serious confusion we must carefully distinguish between offering an explanation for some fact and providing grounds for believing it to be the case. Such confusion is fostered by the fact that the word "why" frequently occurs in two distinct types of locutions, namely, "Why did X occur?" and "Why should one believe that X occurred?" As an example of the first type, we might ask why Marilyn Monroe died. An answer to this *explanation-seeking* why-question is that she took an overdose of sleeping pills. A full explanation would, of course, identify the particular drug and describe its physiological effects. As an example of the second type, we might ask why we believe that she died. The answer to this

evidence-seeking why-question, for me at least, is that it was widely reported in the press. Similarly, to take a more scientific example, it is generally believed by cosmologists that the distant galaxies are receding from us at high velocities. The main *evidence* for this hypothesis is the fact that the light from these galaxies is shifted toward the red end of the spectrum, but this red-shift does not *explain* why the galaxies are traveling away from us. The recession of the galaxies is explained on the basis of the "big bang"—the primordial explosion that sent everything flying off in different directions—*not* by the red shift.

It might be supposed that a confusion of evidential facts with explanatory facts is unlikely to arise, but this supposition would be erroneous. In recent years there has been quite a bit of discussion of the so-called *anthropic principle*. According to certain versions of this principle, earlier states of the universe can be explained by the fact that they involved necessary conditions for the later occurrence of life—particularly human life—as we know it. For example, there must have been stars capable of synthesizing nuclei as complex as carbon. It is one thing to infer, from the undisputed fact that human life exists and would be impossible without carbon, that there is some mechanism of carbon synthesis from hydrogen and helium. It is quite another to claim that the existence of human life at present *explains why* carbon was synthesized in stars in our galaxy.[3]

Another fact that sometimes tends to foster the same confusion is the structural similarity of Hempel's well-known *deductive-nomological (D-N)* model of *scientific explanation* (to be discussed in detail in the next section) and the traditional *hypothetico-deductive (H-D)* schema for *scientific confirmation*. It must be kept in mind, however, that the fundamental aims of these two schemas are quite distinct. We *use* well-confirmed scientific hypotheses, laws, or theories to explain various phenomena. The idea behind deductive-nomological explanation is that, given the truth of all of the statements involved—both those that formulate the explanatory facts and the one that asserts the occurrence of the fact-to-be-explained—the logical relation between premises and conclusion shows that the former explain why the latter obtained. The function of the explanation is not to establish (or support) the truth of its conclusion; that is already presupposed when we accept it as a correct explanation. The idea behind the hypothetico-deductive method, in contrast, is that the given logical schema can be employed to provide evidential support for a hypothesis whose truth is being questioned. The statement that is supposed to be supported by hypothetico-deductive reasoning is not the conclusion in the schema, but rather, one of its premises.[4]

Another, closely related, possible source of confusion is the recent popularity of the slogan "inference to the best explanation." As Gilbert Harman has pointed out, we sometimes use the fact that a certain statement, if true, would explain something that has happened as evidence for the truth of that statement (Harman 1965). A detective, attempting to solve a murder, may consider the possible explanations of the crime, and infer that the 'best' one is true. To describe what is going on here it will be useful to appeal to a distinction (made by Hempel and Oppenheim) between potential explanations and actual explanations. A potential explanation has all of the characteristics of a correct–i.e., actual–explanation, except possibly for the truth of the premises. Harman maintains that we canvass the available potential explanations and infer that the 'best' of these is the actual explanation. As in the case of hypothetico-deductive inference, this kind of inference supports the premises of an explanatory argument, not its conclusion, whose truth is taken for granted from the outset. Given the fact that the whole point of the present essay is to discuss a wide variety of views on the nature of scientific explanation, we are hardly in a position at this stage of our investigation to say much of anything about what constitutes 'the best explanation.' And application of this principle of inference obviously presupposes some explication of *explanation*.

0.2 The Received View

Our story begins in 1948 with the publication of the above-mentioned classic article, "Studies in the Logic of Explanation," by Hempel and Oppenheim. This landmark essay provides the initial document of the old consensus concerning the nature of scientific explanation that emerged around the middle of the twentieth century. It is the fountainhead from which the vast bulk of subsequent philosophical work on scientific explanation has flowed–directly or indirectly.

According to that account, a D-N explanation of a particular event is a valid deductive argument whose conclusion states that the event to be explained did occur. This conclusion is known as the *explanandum-statement*. Its premises–known collectively as the *explanans*–must include a statement of at least one general law that is essential to the validity of the argument–that is, if that premise were deleted and *no other change* were made in the argument, it would no longer be valid. The explanation is said to subsume the fact to be explained under these laws; hence, it is often called "the covering law model." An argument fulfilling the foregoing conditions qualifies as a *potential explanation*. If, in addition, the statements constituting the explanans are true, the argument qualifies as a *true explanation* or simply an *explanation* (of the D-N type).

From the beginning, however, Hempel and Oppenheim (1948, 250–51) recognized that not all legitimate scientific explanations are of the D-N variety; some are probabilistic or statistical. In "Deductive-Nomological vs. Statistical Explanation" (1962) Hempel offered his first account of statistical explanation; to the best of my knowledge this is the first attempt by any philosopher to give a systematic characterization of probabilistic or statistical explanation.[5] In "Aspects of Scientific Explanation" (1965) he provided an improved treatment. This account includes two types of statistical explanation. The first of these, the *inductive-statistical (I-S)*, explains particular occurrences by subsuming them under statistical laws, much as D-N explanations subsume particular events under universal laws. There is, however, a crucial difference: D-N explanations subsume the events to be explained deductively, while I-S explanations subsume them inductively. An explanation of either kind can be described as *an argument to the effect that the event to be explained was to be expected by virtue of certain explanatory facts*. In a D-N explanation, the event to be explained is deductively certain, given the explanatory facts (including the laws); in an I-S explanation the event to be explained has high inductive probability relative to the explanatory facts (including the laws).

On Hempel's theory, it is possible to explain not only particular events but also general regularities. Within the D-N model, universal generalizations are explained by deduction from more comprehensive universal generalizations. In the second type of statistical explanation, the *deductive-statistical (D-S)*, statistical regularities are explained by deduction from more comprehensive statistical laws. This type of statistical explanation is best regarded as a subclass of D-N explanation.

Table 1 shows the four categories of scientific explanations recognized by Hempel in "Aspects." However, in their explication of D-N explanation in 1948, Hempel and Oppenheim restrict their attention to explanations of particular facts, and do not attempt to provide any explication of explanations of general regularities. The reason for this restriction is given in the notorious footnote 33:

> The precise rational reconstruction of explanation as applied to general regularities presents peculiar problems for which we can offer no solution at present. The core of the difficulty can be indicated by reference to an example: Kepler's laws, *K*, may be conjoined with Boyle's law, *B*, to [form] a stronger law *K.B*; but derivation of *K* from the latter would not be considered an explanation of the regularities stated in Kepler's laws; rather, it would be viewed as representing, in effect, a pointless "explanation" of Kepler's laws by themselves. The derivation of Kepler's laws from Newton's laws of motion and

Explananda / Laws	Particular Facts	General Regularities
Universal Laws	D-N Deductive-Nomological	D-N Deductive-Nomological
Statistical Laws	I-S Inductive-Statistical	D-S Deductive-Statistical

Table 1

gravitation, on the other hand, would be recognized as a genuine explanation in terms of more comprehensive regularities, or so-called higher-level laws. The problem therefore arises of setting up clear-cut criteria for the distinction of levels of explanation or for a comparison of generalized sentences as to their comprehensiveness. The establishment of adequate criteria for this purpose is as yet an open problem. (Hempel and Oppenheim 1948, 273; future citations, H-O 1948)

This problem is not resolved in any of Hempel's subsequent writings, including "Aspects of Scientific Explanation."

Chapter XI of Braithwaite's *Scientific Explanation* is entitled "Explanation of Scientific Laws," but it, too, fails to address the problem stated in the Hempel-Oppenheim footnote. Indeed, on the second page of that chapter Braithwaite says,

> To explain a law is to exhibit an established set of hypotheses from which the law follows. It is not necessary for these higher-level hypotheses to be established independently of the law which they explain; all that is required for them to provide an explanation is that they should be regarded as established and that the law should follow logically from them. It is scarcely too much to say that this is the whole truth about the explanation of scientific laws . . . (Braithwaite 1953, 343)

It would appear that Braithwaite is prepared to say that the deduction of Kepler's laws from the conjunction of Kepler's laws and Boyle's law – or the conjunction of Kepler's laws and the law of diminishing marginal utility of money (if you accept the latter as an established law) – is a bona fide explanation of Kepler's laws. However, inasmuch as Braithwaite's book does not contain any citation of the Hempel-Oppenheim paper, it may be that he was simply unaware of the difficulty, at least in this precise form. This problem was addressed by Michael Friedman (1974); we shall discuss his seminal article in §3.5 below. It was also treated by John Watkins (1984); his approach will be discussed in §4.10. Since the same problem obviously applies to D-S explanations, it affects both sectors in the right-hand column of Table 1.

The 1948 Hempel-Oppenheim article marks the division between the prehistory and the history of modern discussions of scientific explanation.[6] Hempel's 1965 "Aspects" article is the central document in the hegemony (with respect to scientific explanation) of logical empiricism, which held sway during roughly the third quarter of the present century. Indeed, I shall use the phrase *the received view* to refer to accounts similar to that given by Hempel in "Aspects." According to the received view, I take it, every legitimate scientific explanation belongs to one of the four sectors of Table 1. As we have seen, the claim of the received view to a comprehensive theory of scientific explanation carries a large promissory note regarding explanations of laws.

The First Decade (1948–57)
Peace in the Valley
(but Some Trouble in the Foothills)

With hindsight we can appreciate the epoch-making significance of the 1948 Hempel-Oppenheim paper; as we analyze it in detail we shall see the basis of its fertility. Nevertheless, during the first decade after its appearance it had rather little influence on philosophical discussions of explanation. To the best of my knowledge only one major critical article appeared, and it came at the very end of the decade (Scheffler 1957); it was more a harbinger of the second decade than a representative of the first. Indeed, during this period not a great deal was published on the nature of scientific explanation in general (in contrast to explanation in particular disciplines).

Braithwaite's (1953) might come to mind as a possible major exception, but we should not be misled by the title. In fact this book contains hardly any explicit discussion of the topic. Braithwaite remarks at the outset that "to understand the way in which a science works, and the way in which it provides explanations of the facts which it investigates, it is necessary to understand the nature of scientific laws, and what it is to establish them" (1953, 2). He then proceeds to discuss at length the nature of scientific deductive systems, including those that involve statistical laws, as well as those that involve only universal laws. Throughout this detailed and illuminating discussion he seems to be assuming implicitly that scientific explanation consists in somehow embedding that which is to be explained in such a deductive system. In adopting this view he appears to be anticipating the Friedman-Kitcher global unification approach, which will be discussed in §3.5 below. However, he has little to say explicitly about the relationship between deductive systems and scientific explanation.[1]

The final two chapters take up some specific issues regarding scientific explanation, and in the course of these chapters Braithwaite makes a few general remarks in passing. For example, the penultimate chapter opens with the statement, "Any proper answer to a 'Why?' question may be said to be an explanation of a sort" (319). In the final chapter he remarks, similarly, "an explanation, as I understand the use of the word, is an answer to a 'Why?' question which gives some intellectual satisfaction" (348–49). These comments are, without doubt, intended for construal in terms of his foregoing discussion of formal systems, but he does not spell out the connections. In the absence of explicit analyses of the nature of why questions, of what constitutes a "proper answer," or of the notion of "intellectual satisfaction," such passing remarks, however suggestive, leave much to be desired. The fact that Braithwaite's book nowhere cites the Hempel-Oppenheim article is eloquent testimony to the neglect of that essay during the first decade.

During this decade interest focused chiefly on two sets of special issues that had been sparked by earlier work. One of these concerned the nature of historical explanation, and the question of whether historical explanations must involve, at least implicitly, appeals to general laws. Much of this discussion took as its point of departure an earlier paper by Hempel (1942). The other dealt with the question of teleological or functional explanation; it came out of the longstanding controversy over mechanism vs. teleology (see H-O 1948, §4). On this specific issue, as we shall see, Braithwaite's book does provide significant contributions. We shall return to these special topics in §1.2 and §1.3, respectively, and we shall find an important connection between them.

1.1 The Fountainhead: The Deductive-Nomological Model

The 1948 Hempel-Oppenheim paper makes no pretense of explicating anything other than D-N explanations of particular occurrences—represented by the upper left-hand sector of Table 1. It will be useful to look in some detail at their treatment of this case. We must distinguish, in the first place, between the general conditions of adequacy for any account of this type of explanation, as laid down in Part I, and the actual explication spelled out in Part III.

The general conditions of adequacy are divided into two groups, logical and empirical. Among the logical conditions we find

(1) the explanation must be a valid deductive argument,
(2) the explanans must contain essentially at least one general law,
(3) the explanans must have empirical content.

The only empirical condition is:

(4) the sentences constituting the explanans must be true.

Although these criteria may seem simple and straightforward, they have been called into serious question. We shall return to this matter a little later.

The general notion of D-N explanation can be represented in the following schema offered by Hempel and Oppenheim, where the arrow signifies deductive entailment. It should also be noted that these criteria of adequacy are meant to apply to D-N explanations of general regularities even though Hempel and Oppenheim do not attempt to provide an explicit explication of explanations of this type. Since the derivation of a narrower generalization (e.g., the behavior of double stars) from a more comprehensive theory (e.g., celestial mechanics) does not

$C_1, C_2, \ldots, C_k$	Statements of antecedent conditions	Explanans
$L_1, L_2, \ldots, L_r$	General laws	
E	Description of the empirical phenomenon to be explained	Explanandum

require any antecedent conditions, they deliberately refrain from requiring that the explanans contain any statements of antecedent conditions.

One of the most vexing problems arising in this context is the characterization of law-sentences. It obviously has crucial importance for the D-N model, as well as for any covering law conception of scientific explanation. Following a strategy introduced by Nelson Goodman (1947), Hempel and Oppenheim (1948, 264–70) attempt to define the broader notion of a *lawlike sentence*. Only true sentences are classified as law-sentences; lawlike sentences have all the characteristics of law-sentences, with the possible exception of truth. Thus every law-sentence is a lawlike sentence, but not all lawlike sentences are laws. Informally, lawlike sentences have four properties:

(1) they have universal form,
(2) their scope is unlimited,
(3) they do not contain designations of particular objects, and
(4) they contain only purely qualitative predicates.

Let us consider the reasons for requiring these characteristics. With regard to (1) and (2) it is intuitively plausible to expect laws of nature to be general laws whose variables range over the entire universe. Newton's laws of universal gravitation and motion apply to all bodies in the universe, and their scope is not restricted in any way. These are paradigms of lawlike statements. However, an apparently universal statement, such as "All Apache pottery is made by women," would not qualify as lawlike because its scope is restricted. Likewise, the statement, "All living things contain water," if tacitly construed to be restricted to living things on earth, would not qualify as lawlike. In contrast, however, "All pure gold is malleable"—though it may appear to have a scope limited to golden objects—is nevertheless a universal generalization of unlimited scope, for it says of each object in the universe that, if it consists of gold, it is malleable. The distinction among the foregoing examples between those that qualify as lawlike and those that do not relates to characteristic (3). The statement about Apache pottery makes explicit reference to a particular group of people, the Apache. The statement about living things, if construed as suggested, refers implicitly to our particular planet.[2]

Why does it matter, with respect to lawlikeness, whether a statement refers to a particular of some sort—a particular time, place, object, person, group, or nation? Consider a simple example. Suppose it happens to be true (because I like golden delicious apples) that all of the apples in my refrigerator are yellow. This statement involves reference to a particular person (me), a particular thing (my refrigerator), and a particular time (now). Even given my taste in apples it is not impossible for my refrigerator to contain apples of different colors. Moreover, there is no presumption that a red delicious apple would turn yellow if it were placed in my refrigerator.

The problem that arises in this context is to distinguish between laws and *accidental generalizations*. This is a crucial issue, for laws have explanatory force, while accidental generalizations, *even if they are true*, do not. It obviously is no explanation of the color of an apple that it happens to reside in my refrigerator at some particular time.

If a statement is to express a law of nature it must be true. The question is, what characteristics, in addition to truth, must it possess? *Generality* is one such characteristic: laws must apply universally and they must not contain special provisions or exceptions for particular individuals or groups. The *ability to support counterfactuals* is another: they must tell us what would happen if. . . . If this table salt were placed in water, it would dissolve. If this switch were closed, a current would flow in this circuit.[3] *Modal import* is another: laws delineate what is necessary, possible, or impossible. We are not talking about logical modalities, of course; we are concerned with what is *physically* necessary, possible, or impossible. According to relativity theory it is physically impossible to send a signal faster than light *in vacuo*; according to the first law of thermodynamics it is physically impossible to construct a perpetual motion machine (of the first type). Accidental generalizations, even if true, do not support counterfactuals or possess modal import.

Even if a given statement does not contain explicit designations of particular objects, it may involve implicit reference to one or more particulars. Such references may be hidden in the predicates we use. Terms like "lunar," "solar," "precolumbian," and "arctic," are obvious examples. Because such terms refer to particulars they do not qualify as purely qualitative. By stipulating, in property (4) above, that laws contain only purely qualitative predicates, this sort of implicit reference to particulars, is excluded. Properties (3) and (4) are designed to rule out as accidental those universal generalizations that contain either explicit or implicit reference to particulars.

As Hempel and Oppenheim are fully aware, the prohibition against reference to particulars they impose is extremely stringent. Under that restriction, neither Galileo's law of falling bodies (which refers explicitly to the earth) nor Kepler's laws of planetary motion (which refer explicitly to our solar system) would qualify as laws or lawlike statements. As we shall see, because of this consideration they distinguish between fundamental and derived laws. The foregoing restrictions apply only to the fundamental laws. Any universal statement that can be deduced from fundamental laws qualifies as a derived law.

Yet, in spite of their careful attention to the problem of distinguishing between lawful and accidental generalizations, Hempel and Oppenheim did not succeed in explicating that distinction. Consider the following two statements:

(i) No signal travels faster than light.
(ii) No gold sphere has a mass greater than 100,000 kg.

Let us suppose, for the sake of argument, that both are true. Then we have two true (negative) universal generalizations. Both have universal form. Neither is restricted in scope; they refer, respectively, to signals and gold spheres anywhere in the universe at any time in its history—past, present, or future. Neither makes explicit reference to any particulars. Both statements satisfy characteristics (1)-(3). One might argue that the predicate "having mass greater than 100,000 kg" is not purely qualitative, since it contains a reference to a particular object—namely, the international prototype kilogram. But this difficulty can be avoided by expressing the mass in terms of atomic mass units (which refer, not to any particular object, but to carbon-12 atoms in general). Thus, with (ii) suitably reformulated, we have two statements that satisfy characteristics (1)–(4), one of which seems patently lawful, the other of which seems patently accidental. The contrast can be heightened by considering

(iii) No enriched uranium sphere has a mass greater than 100,000 kg.

Since the critical mass for enriched uranium is just a few kilograms, (iii) must be considered lawful.

Both statements (i) and (iii) have modal import, whereas (ii) does not. It is physically impossible to send a message faster than light and it is physically impossible to fabricate an enriched uranium sphere of mass greater than 100,000 kg. It is not physically impossible to fabricate a gold sphere of mass greater than 100,000 kg.[4] Likewise, statements (i) and (iii) support counterfactuals, whereas (ii) does not. If something were to travel faster than light it would not transmit information.[5] If something were a sphere with mass greater than 100,000 kg it would not be composed of enriched uranium. In contrast, we cannot legitimately conclude from the truth of (ii) that if something were a sphere with mass greater than 100,000 kg, it would not be composed of gold. We cannot conclude that if two golden hemispheres with masses greater than 50,000 kg each were brought together, they would explode, suffer gravitational collapse, undergo severe distortion of shape, or whatever, instead of forming a sphere.

Lawfulness, modal import, and support of counterfactuals seem to have a common extension; statements either possess all three or lack all three. But it is extraordinarily difficult to find criteria to separate those statements that do from those that do not. The three characteristics form a tight little circle. If we knew which statements are lawful,we could determine which statements have modal import and support counterfactuals. But the way to determine whether a statement has modal import is to determine whether it is a law. The same consideration applies to support of counterfactuals; to determine which statements support counterfactuals we need to ascertain which are laws.[6] The circle seems unbroken. To determine to which statements any one of these characteristics applies we need to be able to determine to which statements another of them applies.

There are, of course, a number of differences between statements (i) and (iii) on the one hand and statement (ii) on the other. For example, I am much less confident of the truth of (ii) than I am of (i) or (iii). But this is a psychological statement about my state of belief. However, we are assuming the truth of all three statements. Given that all three are true, is there any objective difference in their status, or is the sole difference psychological? Again, (i) and (iii) fit closely with a well-integrated body of physical theory, while (ii) does not.[7] But given that all three are true, is this more than an epistemic difference?[8] Further, there are differences in the ways I might come to know the truth of (ii), as opposed to coming to know the truth of (i) and (iii). But is this more than an epistemic or psychological difference? Still further, I would much more readily give up my belief in (ii) than I would my belief in (i) or (iii).[9] But is this more than a pragmatic difference? The unresolved question is this: is there any objective distinction between laws and true accidental generalizations?[10] Or is the distinction wholly psychological, epistemic, or pragmatic?

In his 1953 book, Braithwaite places considerable emphasis upon the nature of laws and their place in science. He writes, "In common with most of the scientists who have written on philosophy of science from Ernst Mach and Karl Pearson to Harold Jeffreys, I agree with the principal part of Hume's thesis – the part asserting that universals of law are objectively just universals of fact, and that in nature there is no extra element of necessary connexion" (1953, 294). In chapter IX he defends the view that "the difference between universals of law and universals of fact [lies] in the different roles they play in our thinking rather than in any difference in their objective content" (294–95).[11]

The most ambitious attempt by any of the logical empiricists to deal with these problems concerning the nature of laws was given by Hans Reichenbach (1954),[12] the year just after the publication of Braithwaite's book. It had been anticipated by his discussion of the same topics in his symbolic logic book (1947, chap. VIII). Reichenbach's very first requirement on law-statements makes the distinction between laws and accidental generalizations an epistemic one, for it refers explicitly to the types of evidence by which such statements are supported. It should be remarked, incidentally, that Reichenbach was not addressing these problems in the context of theories of scientific explanation.

The problem of characterizing law-statements is one that has not gone away. Skipping ahead to subsequent decades, we may note that Ernest Nagel's magnum opus on scientific explanation, published near the beginning of the second decade, has a sensitive and detailed discussion of this problem, but one that remains inconclusive.[13] Around the beginning of the third decade, Nicholas Rescher's book *Scientific Explanation* offers an extended discussion which concludes that lawfulness does not reflect objective factors in the world, but rather rests upon our imputations, and is consequently mind-dependent (1970, 97–121; see also Rescher 1969). In the fourth decade, to mention just one example among many, Brian Skyrms (1980) offers a pragmatic analysis. The fifth decade will see the publication of an extremely important work on the subject, *Laws and Symmetries*, by Bas van Fraassen. But let us return to the first decade.

To carry out their precise explication, Hempel and Oppenheim introduce a formal language in which scientific explanations are supposed to be formulated. It is a standard first order functional calculus without identity, but no open sentences are allowed. All individual variables are quantified, so generality is always expressed by means of quantifiers. Two semantical conditions are imposed on the interpretation of this language: First, the range of the individual variables consists of all physical objects in the universe or of all spatio-temporal locations; this ensures that requirement (2) on lawlike statements – that their scope be unlimited – will be fulfilled, for there is no limit on the range of the variables that are universally (or existentially) quantified. Second, the primitive predicates are all purely qualitative; this feature of the interpretation of the language is, of course, a direct reflection of the fourth requirement on lawlike statements. The explication of D-N explanation of particular occurrences is given wholly in semantical terms.

Before going into the details of the formal language, we must acknowledge a fundamental problem regarding the second of the foregoing semantical conditions, namely, the concept of a purely qualitative predicate. In his well-known book *Fact, Fiction, and Forecast* (1955), Nelson Goodman poses what he calls "the new riddle of induction" in terms of two predicates, "grue" and "bleen," that he constructs for that purpose. Select quite arbitrarily some future time *t* (say the beginning of the twenty-first century). "The predicate 'grue' applies to all things examined before *t* just in case they are green but to other things just in case they are blue" (1955, 74). "Bleen" applies to things examined before *t* just in case they are blue but to other things just in case they are green (1955, 79). The question Goodman poses is whether we should inductively project that twenty-first century emeralds will be green or that they will be grue. The same problem had originally been posed by Goodman in 1947.

In an answer to Goodman's query, Carnap maintained that "grue" and "bleen," in contrast to "blue" and "green," are not purely qualitative predicates, because of the reference to a particular time in their definitions. He proposes to resolve Goodman's problem by restricting the predicates of his languages for confirmation theory to purely qualitative ones (Carnap 1947).[14] Goodman demurs:

> . . . the argument that the former but not the latter are purely qualitative seems to me quite unsound. True enough, if we start with "blue" and "green," then "grue" and "bleen" will be explained in terms of "blue" and "green" and a temporal term. But equally truly, if we start with "grue" and "bleen," then "blue" and "green" will be explained in terms of "grue" and "bleen" and a temporal term; "green," for example, applies to emeralds examined before time *t* just in case they are grue, and to other emeralds just in case they are bleen. Thus qualitativeness is an entirely relative matter.(1947)

It is now generally conceded that Carnap's attempt to characterize purely qualitative predicates was inadequate to deal with the problem Goodman raised. Many philosophers (including this one – (W. Salmon 1963)) have tried to make good on the distinction Carnap obviously had in mind. Whether any of these other efforts have been successful is a matter of some controversy; at any rate, no particular solution has gained general acceptance.

Our discussion, so far, has been largely preparatory with respect to the official Hempel-Oppenheim explication. We may now return to the formal language offered by Hempel and Oppenheim. Several different types of sentences must be distinguished. To begin, an *atomic sentence* is one that contains no quantifiers, no variables, and no sentential connectives. It is a sentence that attributes a particular property to a given individual (e.g., "George is tall") or asserts that a particular relation holds among two or more given individuals (e.g., "John loves Mary"). A *basic sentence* is either an atomic sentence or the negation of an atomic sentence; a basic sentence contains no quantifiers, no variables, and no binary sentential connectives. *Singular* (or *molecular*) *sentences* contain no quantifiers or variables, but they may contain binary sentential connectives (e.g., "Mary loves John or Mary loves Peter"). A *generalized sentence* contains one or more quantifiers followed by an expression containing no quantifiers (e.g., "All humans are mortal"). Since any sentence in first order logic can be transformed into *prenex normal form*, any sentence containing quantifiers can be written as a generalized sentence. *Universal sentences* are generalized sentences containing only universal quantifiers. A generalized (universal) sentence is *purely generalized (universal) if it contains no proper names of individuals. A generalized (universal)* sentence is an *essentially generalized (universal)* sentence that is not equivalent to any singular sentence. With these definitions in hand we can proceed to explicate the fundamental concepts involved in scientific explanation.

The first concept with which we must come to terms is that of a law of nature, and, as we have seen, it is one of the most problematic. Hempel and Oppenheim distinguish between *lawlike sentences* and genuine laws, and also between *fundamental* and *derivative* laws. The following series of definitions is offered:[15]

(7.3a) A *fundamental lawlike sentence* is any purely universal sentence; a *fundamental law* is purely universal and true.

(7.3b) A *derivative law* is a sentence that is essentially, but not purely, universal and is deducible from some set of fundamental laws.

(7.3c) A *law* is any sentence that is either a fundamental or a derived law.

We have already canvassed the fundamental problems encountered in this characterization of laws.

Interestingly, the concept of law does not enter into the formal explication of D-N explanation; instead, the notion of *theory* is employed.

(7.4a) A *fundamental theory* is any sentence that is purely generalized and true.

(7.4b) A *derivative theory* is any sentence that is essentially, but not purely, generalized and is derivable from fundamental theories.

(7.4c) A *theory* is any fundamental or derivative theory.

Note that the concept of a theory-like sentence is not introduced.

According to the foregoing definitions, every law is a theory and every theory is true. As the term "theory" is used in this context, there is no presumption that theories refer to unobservable entities, or that they involve any sort of special theoretical vocabulary. The difference between laws and theories is simply that theories may contain existential quantifiers, while laws contain only universal quantifiers. Clearly, many of the scientific laws or theories that are employed in explanation contain existential quantifiers. To say, for example, that every comet has a tail, that every atom has a nucleus, or that every mammal has a heart, involves a universal quantifier followed by an existential quantifier – i.e., for every x there is a y such that . . . Hempel and Oppenheim say nothing about the order in which quantifiers must occur in theories. That leaves open the interesting question of whether explanatory theories may have existential quantifiers preceding all of the universal quantifiers, or whether explanatory theories need contain any universal quantifiers at all.[16] It is perhaps worth explicit mention in this context that universality and generality are not coextensive. Existentially quantified statements are general in the sense that they involve variables having the universe as their range. To say, "there exists an x such that . . . " means that within the whole domain over which x ranges there is at least one object such that. . . . Such statements have generality without being universal. The question remains whether universality is a necessary requirement for explanatory theories, or whether generality is sufficient. As we shall see in connection with the next set of formal definitions, Hempel and Oppenheim are willing to settle for the latter alternative.

We have finally arrived at the stage at which Hempel and Oppenheim offer their formal explication of scientific explanation. The concept of a potential explanation comes first:

(7.5) $<T,C>$ is a potential explanans of E (a singular sentence) *only if*
(1) T is essentially general and C is singular, and
(2) E is derivable from T and C jointly, but not from C alone.

It would be natural to suppose that (7.5) would constitute a definition of "potential explanans," but Hempel and Oppenheim are careful to point out that it provides only a necessary condition. If it were taken as sufficient as well, it would leave open the possibility that "any given particular fact could be explained by means of any true lawlike sentence whatever" (H-O 1948, 276). They offer the following example. Let the explanandum-statement E be "Mount Everest is snowcapped" and let the theory T be "All metals are good conductors of heat." Take a singular sentence T_s that is an instance of T–e.g., "If the Eiffel Tower is metal it is a good conductor of heat." Now take as the singular sentence C the sentence T_s implies E–i.e., "If the fact that the Eiffel Tower is made of metal implies that it is a good conductor of heat, then Mount Everest is snowcapped." Because E is true, C must be true, for C is a material conditional statement with a true consequent. Thus,

$\vdash C = T_s \supset E$	definition
$T_s \supset E$	assumption
$\vdash T \supset T_s$	instantiation
$T \supset E$	hypothetical syllogism
$\vdash C \supset (T \supset E)$	conditional proof
$\vdash C\,.\,T \supset E$	tautology

It is evident that C does not, by itself, entail E. Therefore $<T,C>$ satisfies (7.5). But it is manifestly absurd to claim that the law about metals being good conductors of heat is the key law in the explanation of snow on Mount Everest.

The obvious difficulty with this example is that C's truth can be fully certified only on the basis of the truth of E. Evidently, some restriction must be placed on the singular sentence C that is to serve as the statement of antecedent conditions in the explanans. If knowing that the explanandum-statement is true is the only way to establish the truth of C, then in some important sense, in appealing to C, we are simply using E to explain E. Indeed, given that T is true, there must be some way to establish the truth of C without appealing to E. Hempel and Oppenheim formulate the needed restriction as follows:

(3) T must be compatible with at least one class of basic sentences which has C but not E as a consequence.

That is to say, given that the theory T is true, there must be some way to verify that C is true without also automatically verifying E as well. Adding (3) to the necessary conditions stated in (7.5) gives

(7.8) $<T,C>$ is a potential explanans of E (a singular sentence) *iff*
(1) T is essentially general and C is singular, and

(2) E is derivable from T and C jointly, but not from C alone.
(3) T must be compatible with at least one class of basic sentences which has C but not E as a consequence.[17]

With this definition of "potential explanans" it is a small step to the official explication of "explanans," and hence, "explanation."

(7.6) $<T,C>$ is an explanans of E (a singular sentence) *iff*
(1) $<T,C>$ is a potential explanans of E
(2) T is a theory and C is true.

Taken together, the explanans $<T,C>$ and the explanandum E constitute an explanation of E. This completes the Hempel-Oppenheim explication of *D-N explanation of a particular fact*.

Given the great care with which the foregoing explication was constructed, it would be easy to surmise that it is technically correct. Jumping ahead to the next decade for a moment, we find that such a supposition would be false. As Rolf Eberle, David Kaplan, and Richard Montague (1961) showed (roughly), on the foregoing explication any theory T can explain any fact E, where T and E have no predicates in common, and are therefore, intuitively speaking, utterly irrelevant to one another. Suppose, for example, that T is "(x)Fx" (e.g., "Everyone is imperfect.") and E is "Ha" (e.g., "C. G. Hempel is male.").[18] We can formulate another theory T′ that is a logical consequence of T:

$$T' =_{df} (x)(y)[Fx \vee (Gy \supset Hy)]$$

T′ is of purely universal form, and, on the assumption that T is true, it is true as well. As a singular sentence, take

$$C =_{df} (Fb \vee \sim Ga) \supset Ha$$

For the sake of our concrete interpretation, we can let "Gx" mean "x is a philosopher" and let "b" stand for W. V. Quine. It can now be shown that $<T',C>$ constitutes an explanans of E.

(1)	$(x)(y)[Fx \vee (Gy \supset Hy)]$	premise (T′)
(2)	$(Fb \vee \sim Ga) \supset Ha$	premise (C)
(3)	$(Fb \supset Ha)\,.\,(\sim Ga \supset Ha)$	equivalent to (2)
(4)	$Fb \supset Ha$	simplification (3)
(5)	$\sim Ga \supset Ha$	simplification (3)
(6)	$Fb \vee (Ga \supset Ha)$	instantiation (1)
(7)	$\sim Fb \supset (Ga \supset Ha)$	equivalent to (6)
(8)	$Fb \vee \sim Fb$	tautology
(9)	$Ha \vee (Ga \supset Ha)$	dilemma (4, 7, 8)
(10)	$Ha \vee \sim Ga \vee Ha$	equivalent to (9)
(11)	$\sim Ga \vee Ha$	equivalent to (10)

(12)	$Ga \supset Ha$	equivalent to (11)
(13)	$Ga \vee \sim Ga$	tautology
(14)	Ha	dilemma (5, 12,13)

As we have seen, T′ is essentially general, C is singular, and E is derivable from T′ and C. Hence, conditions (1) and (2) of (7.8) are satisfied. Now, consider the set of basic sentences $\{\sim Fb, Ga\}$; obviously it does not entail E (i.e., Ha). But it does entail C, as follows:

(1)	$\sim Fb \,.\, Ga$	premise
(2)	$(\sim Fb \,.\, Ga) \vee Ha$	addition (1)
(3)	$\sim(\sim Fb \,.\, Ga) \supset Ha$	equivalent to (2)
(4)	$(Fb \vee \sim Ga) \supset Ha$	DeMorgan (3)

Thus, condition (3) of (7.8) is also fulfilled; $<T',C>$ is an explanans for E. We should remember that $<T,C>$ has not been shown to be an explanans for E, so it has not been shown that any arbitrarily chosen theory explains any particular fact (as long as they share no predicates). But what has been shown is something like this: from the law of diminishing marginal utility of money (which, for the sake of argument, I take to be true), we can deduce a theory which, together with certain singular statements, provides an explanation of the explosion of the *Challenger* space-shuttle vehicle. It can hardly be doubted that Eberle, Kaplan, and Montague offered a recipe (in fact, several recipes) for constructing counterexamples that are damaging to the Hempel-Oppenheim explication.

To assess the nature of the damage, let us go back to Hempel and Oppenheim's set of necessary conditions (7.5), which at first blush looks like a suitable explication of *potential D-N explanans*. Had they offered it as such, their explication would have been technically defective, for it would have been vulnerable to the Mount Everest counterexample. However, it would not have posed a profound problem for their enterprise, since the technical defect could have been repaired by technical tinkering, namely, by adding the third condition of (7.8). Fortunately, they anticipated the problem and found a solution before their essay was published.[19]

The same attitude should, I believe, be taken to the problem discovered 13 years later by Eberle, Kaplan, and Montague. They pointed out a genuine technical defect, but again, it was one that could be repaired by technical tinkering. Indeed, in the same volume of the same journal, Kaplan (1961) provided one way of making the repair. Shortly thereafter another was offered by Jaegwon Kim (1963).

It is worth noting, I believe, that both the Mount Everest example and the Eberle-Kaplan-Montague type examples exploit a well-known feature of standard truth-functional logic. This feature is the principle of addition that allows one to deduce "$p \vee q$" from "p"—where "q" can be any arbitrary sentence whatever. This

principle is closely related to one of the so-called paradoxes of material implication, namely, the fact that if "p" is true then it is materially implied by any arbitrary sentence whatever—e.g., "$\sim q$"—so that, given "p" we can infer "$\sim q \supset p$" (which is equivalent to "$p \vee q$"). In the Mount Everest example, we chose C as $T_s \supset E$, which we were prepared to assert on the basis of the truth of E. In the Eberle-Kaplan-Montague example we chose C as "$(Fb \vee \sim Ga) \supset Ha$"—where "Ha" is E. In addition, from the theory "$(x)Fx$" we derived "$(x)(y)[Fx \vee (Gy \supset Hy)]$" which is equivalent to "$(x)Fx \vee (y)(Gy \supset Hy)$." Clearly the source of difficulty is that these moves are precisely the sorts that allow the introduction of irrelevancies—they are the very principles that are excluded in relevance logics. The technical problem with the Hempel-Oppenheim explication is simply to find ways of blocking them in the context of explanation. Kaplan and Kim have shown how to do that. Kim's revision consists in adding a further requirement to those already included in (7.8):

(4) E must not entail any conjunct of the conjunctive normal form of C.[20]

In 1965 Hempel could be quite sanguine about the technical details of the corrected explication (Hempel 1965, 294–95).

It should be emphatically noted that the official explication—as given in the augmented definition (7.8) is very different from the set of requirements of adequacy given in the first section of the Hempel and Oppenheim essay. This is to be expected. The informal conditions of adequacy are part of the clarification of the explicandum; the explication is the formal definition of the improved concept that is to replace the original vague concept of explanation. What is perhaps surprising is the complexity of the formal explication.

Now that we have looked at the nitty-gritty details of the Hempel-Oppenheim explication, let us return to a consideration of the fundamental philosophical issues to which it gives rise. Where does this discussion leave us? First, there is a two-part agenda, explicitly stated in that essay, of items that are needed to complete the received view:

(1) Explications of one or more models of probabilistic or statistical explanation.

(2) An explication of D-N explanation of laws.

Second, there are, as we have seen, two serious lacunae in the Hempel-Oppenheim explication of D-N explanation of particular facts:

(3) A satisfactory explication of the concept of a purely qualitative predicate.

(4) A satisfactory explication of the concept of a law of nature.

Third, as we shall see in greater detail when we discuss the second decade, the Hempel-Oppenheim essay advances several important philosophical theses that have been the subject of much controversy:

(5) The inferential conception of scientific explanation–i.e., the thesis that all legitimate scientific explanations are arguments of one sort or another. This thesis is involved in Hempel and Oppenheim's first condition of adequacy for scientific explanations. A number of subsequent authors including Michael Scriven, Richard Jeffrey, Bas van Fraassen, Peter Railton, and I–have rejected this inferential conception.

(6) The covering law conception of scientific explanation[21]–i.e., the view that every fully articulated adequate scientific explanation contains one or more laws in an essential fashion. This thesis is set forth in Hempel and Oppenheim's second condition of adequacy. As we shall see, this view is also rejected by a number of authors–including Scriven and van Fraassen. Thomas S. Kuhn strenuously denies it.[22] So do many others who are primarily concerned with the nature of explanation in history.

(7) The explanation/prediction symmetry thesis. According to the strong form of this thesis, as set forth in the Hempel-Oppenheim article, any correct D-N explanation could serve, in appropriate circumstances, as a scientific prediction; conversely, any deductive scientific prediction could, in appropriate circumstances, serve as a D-N explanation. This thesis, which Hempel later extended to I-S explanation as well, has been widely disputed. Hanson, Israel Scheffler, Scriven, van Fraassen, and I–among many others–have been unwilling to accept it.

(8) The role of causality in scientific explanation. Hempel and Oppenheim casually identify causal explanation with D-N explanation (H-O 1948, 250), but their official explication makes no reference to any causal requirements. In "Aspects," Hempel explicitly rejects the idea that causality plays any essential explanatory role (1965, 352). The question is whether to follow Hempel in "Aspects" in expunging causality from explanation altogether, or to find a place for causality, as Scriven, Railton, James H. Fetzer, Paul Humphreys, and I–among many others–have urged. As we shall see, this issue has been the subject of considerable subsequent discussion.

(9) Literal truth of the explanans. This is Hempel and Oppenheim's fourth condition of adequacy. Many authors have maintained that this requirement is too strong–that it should be replaced by a requirement of high degree of confirmation, or by a requirement of approximate truth. For various reasons this requirement has been rejected by Nancy Cartwright, Kuhn, and Larry Laudan.

(10) The possibility of quasi-formal models of explanation. The early logical empiricists seemed confident that many fundamental methodological concepts–such as confirmation and explanation–are amenable to formal semantical explication, more or less as attempted in the Hempel-Oppenheim paper. It now seems to many–including Peter Achinstein and me–that explications of that sort are bound to turn out to be inadequate. Even Hempel in his later writings on scientific explanation eschews the degree of formality employed in the Hempel-Oppenheim explication.

Our review of the essay by Hempel and Oppenheim leaves no doubt whatever about its richness as a source of material for philosophical discussions of the nature of scientific explanation in the second half of the twentieth century. Can anyone seriously dispute the propriety of its designation as the fountainhead?

1.2 Explanation in History and Prehistory

In the decade following the publication of the Hempel-Oppenheim paper, as I pointed out above, not a great deal appeared in print on general problems concerning the nature of scientific explanation.[23] During this time, however, quite a good deal was written on the nature of explanation in history, stimulated, to a large degree, by Hempel's essay, "The Function of General Laws in History" (1942). The focus of much of the discussion was the question of whether historical explanations can or should embody general laws. This literature is directly pertinent to our discussion only if history is considered a science. Since I am not firmly committed to any particular view on this matter, and know very little about history, I shall not pursue the issue here. I refer the reader to Hempel's 1942 article, to §7 of "Aspects," and to the excellent bibliography of that essay.

Whether history is classified as a science or not, there can be no doubt that some sciences have essential historical aspects. Cosmology, geology, and evolutionary biology come immediately to mind. Archaeology (often called "prehistory") is somewhat ambiguous, for it has been pursued, not infrequently, as a humanistic, rather than scientific, discipline.

During the 1950s, an important development occurred in archaeology, namely, the origin of an influential movement called the *New Archaeology*. The main thrust of this movement is the attempt to make archaeology a bona fide science. When Lewis Binford, one of its major founders, was a student at the University of Michigan, Leslie White–a distinguished cultural anthropologist and a teacher of Binford–told him that, to find out what it means to be scientific, he should read some philosophy of science. He did (Binford 1972, 7–8). Although Binford's published work does not contain extensive discussions of philosophical issues, the effect upon him was profound, and he profoundly influenced other archaeologists. The most explicit expression of the effect of philosophy of science upon archaeology was given in Watson, LeBlanc, and Redman's widely used text, *Explanation in Archaeology* (1971, 1984). This book relies heavily on Hempel's work on explanation, especially his *Philosophy of Natural Science* (1966) as well as "Aspects of Scientific Explanation." It is hardly an exaggeration to say that, for these authors, the hallmark of genuine science is the use of hypothetico-deductive confirmation and D-N explanation. They urge their students and colleagues to adopt this kind of methodology. The New Archaeology, which had its heyday in the late 1960s and throughout the 1970s, provides an example of an outstanding influence of philosophy of science upon an actual science. It is still a strong movement. Whether the influence was good or bad is a matter of some controversy. A detailed account of the relationships between the New Archaeology and philosophy of science is provided by Merrilee H. Salmon in *Philosophy and Archaeology* (1982).

1.3 Teleology and Functional Explanation

As we have seen, logical empiricism arose in a philosophical context in which teleology, final causes, purposes, and ends played vital roles. Those scientific philosophers who—like Braithwaite, Hempel, and Nagel—had not abandoned the notion that science can provide legitimate explanations of various kinds of phenomena were deeply concerned to provide accounts that would require admissible scientific explanations to have empirical content. They strove to prohibit pseudo-explanations that appeal to entelechies or final causes.

Nevertheless, there appeared to be cases in the empirical sciences and in common sense of bona fide explanations that are framed in terms of ends or goals. One often explains one's own behavior in this way. Why did I go to the Groceria Italiana? To get fresh pasta. In cases of this sort, however, it is not the future state of procuring pasta that explains my act; it is my preceding desire for the pasta along with the concurrent belief that that particular grocery store was a good place to get it. The explanation would be correct even if I failed to get the pasta because they were all sold out.

Midway through the first decade, Braithwaite addressed the problem of teleological explanation (1953, chap. X).[24] He saw no problem in teleological explanations of actions that are the result of conscious intention, but he also pointed out that there are cases of goal-directed behavior in which *conscious* intent seems to be absent. Among the examples Braithwaite mentions are the behavior of rats in mazes and the operation of homing torpedoes. The basic philosophical problem in these cases, given the absence of conscious intent (or any kind of intent on the part of the torpedo), involves the question of whether a future state can legitimately be said to explain a present fact.

Braithwaite notices two striking features of goal-directed behavior, *plasticity* and *variancy*. Plasticity refers to the fact that, in many cases, the agent has more than one way to achieve the goal; if one means is frustrated another will be tried. An animal trying to arrive at a place where food is located will try different routes if the usual one is blocked. Variancy refers to the fact that the goal can successfully be reached under a wide variety of initial and background conditions. A homing torpedo can—within limits—find its target regardless of the evasive action taken and—again, within limits—regardless of weather or conditions of the sea.

It has been pointed out that discussions of teleological or functional explanation often employ either or both of two models: (1) *goal-directed* behavior, as illustrated by the foregoing examples of the homing torpedo and the rat, and (2) *self-regulating* behavior, as illustrated by a thermostatically controlled heat-pump or maintenance of body temperature by a human or other animal (Canfield 1966, Scheffler 1958). A thermostatically controlled heat-pump, for example, will keep the temperature of a house within a certain temperature range whether the initial temperature is below or above that range by a little or a lot. It will do so under a wide range of outdoor temperatures, wind velocities, conditions of precipitation, etc. The human body also is capable of maintaining an internal body temperature within a narrow normal range under a wide variety of external circumstances.

In some cases, such as the homing torpedo and the heat-pump, we can give a full causal account of the apparently purposeful behavior. In an influential article, Rosenblueth, Wiener, and Bigelow (1943) gave a cybernetic explanation of the behavior of the homing torpedo in terms of negative feedback, and they suggested that the apparently goal-directed behavior of animals can be understood in similar terms. These are Braithwaite's kind of example as well. Ernest Nagel, who dealt mainly with teleological or functional explanation in physiology, focused much more intensively on the *homeostatic* or self-regulating type of example (1956; 1961, chap. 12). Just as we can provide a complete causal account of the way in which the heat-pump controls the house temperature, starting with any set of background conditions within its range of operation, so also, he maintained, could we give a complete causal account of the way in which the human body regulates its internal temperature.

According to Braithwaite, if we give a teleological explanation of the operation of a homing torpedo in terms of the goal of reaching its target, or the heat-pump/thermostat system in terms of the goal of maintaining a given house temperature, that is merely a shorthand for a much more complicated causal explanation, and philosophically it is not very interesting. In other cases, such as the rat in the maze, we may not be able to give a complete causal explanation in physico-chemical terms, without reference to goal-seeking. In that case a teleological explanation is appropriate, legitimate, and interesting. But we have no reason to believe that the causal explanation in purely physico-chemical terms is impossible in principle. Indeed, Nagel argued the stronger point that, in many physiological cases, such causal explanations are already in hand—and that, in such cases, any teleological explanation can, in principle, be replaced by an equivalent causal explanation. Thus, we have no basis for supposing that in any instance of goal-directed behavior is a present fact ultimately explainable only in terms of a future result.

If we stretch the temporal bound just a little beyond the end of the first decade, we can include two other significant contributions to the discussion of teleological or functional explanation—namely, the paper by Israel Scheffler (1958)[25] and a characteristically clear treatment by Hempel (1959). Hempel's discussion brings the D-N model to bear directly on the problem of functional explanation.

Hempel's philosophical concern is rather different from Braithwaite's; he does not focus particularly on the issue of final causation—on the problem of explaining present facts in terms of future goals. Given his later suggestion that subsequent events might sometimes explain earlier events (Hempel 1965, 353–54), that issue might prove embarrassing. He has, instead, a logical point in mind. The kinds of cases in which he is interested are often referred to as *functional explanations*; let us consider some simple examples. In physiology, for instance, the presence of a particular component in a given species of organisms is explained in terms of its ability to perform a function that is indispensable to the continued life or health of organisms of that sort. In humans, the blood contains hemoglobin to convey oxygen from the lungs to the other parts of the body. Without this transport of oxygen a human being could not survive.

In evolutionary biology, for another example, a feature of a species of animal is explained in terms of its enhancement of the chances of survival and reproduction. The large ears of the jackrabbit, which inhabits very hot regions, enable the animal to control its body temperature. When its body temperature rises too high, the animal seeks shade and dilates the many blood vessels in the ears. Blood coming from other parts of the body brings heat which is radiated into the environment, thereby cooling the animal. I have heard that the ears of elephants function in much the same way. In biological evolution adaptation to the environment plays a crucial role, and this appears to involve the attainment of goals.

Anthropology and sociology provide many additional cases of functional explanation. In the study of primitive societies, for instance, such prominent anthropologists as A. R. Radcliffe-Brown (1952) and B. Malinowski (1954) maintain that many—if not all—institutions, customs, and rituals are to be explained on the basis of the social functions they fulfill. The performance of a rain dance, for example, may provide an occasion for social interaction that contributes to the cohesiveness of the society. The influential sociologist R. K. Merton (1950, 1957) also advocates functional analysis in the study of human institutions.[26] He distinguishes carefully between *latent function* and *manifest function*. The rain dance has the manifest function of bringing rainfall, and it may be altogether unsuccessful in fulfilling that function; it has the latent function, however, of promoting social cohesiveness in times of distress, and it may fulfill that function quite effectively. In such cases, the latent function *explains* the survival of a practice that fails miserably to fulfill its manifest function.

Freudian psychoanalysis (which may or may not deserve to be classified as a science) is another rich source of functional explanations. For example, Freud claims that dreams serve as wish-fulfillments to prevent interruption of sleep, which is essential to the health of any human being. He also offers functional explanations of slips of the tongue or pen, and for the presence of neurotic symptoms.

The statements involved in the foregoing accounts—with the possible exception of psychoanalysis—are empirically testable and scientifically legitimate. Even if one denies that psychoanalysis is a genuine science, the basis for that judgment should not be the presence of functional explanations; the crucial issue lies in the empirical testability of its claims. It appears, then, that we have a number of different sciences that make legitimate use of functional explanations. The problem, as far as Hempel is concerned, is that functional explanations do not fit any of the logical patterns for scientific explanation, including those for inductive or statistical explanation, recognized by the received view.[27] It is *very* significant that Hempel entitled his article "The Logic of Functional Analysis" rather than "The Logic of Functional Explanation."

The fundamental problem Hempel faces can easily be seen by schematizing a simple case. He chooses the example

(3.1) The heartbeat in vertebrates has the function of circulating blood through the organism. (1959, 305)

In attempting to understand the meaning of such a statement we might be tempted simply to substitute the word "effect" for the word "function," since the circulation of blood is an effect of the beating of the heart. But the beating of the heart also has other effects that we would be unwilling to consider functions—for example, it has the effect of producing heart sounds, but that is not one of its functions. The general idea is that a function must be important to the health or survival of the organism. To do more justice to the significance of (3.1) he reformulates it as

(3.3) The heartbeat has the effect of circulating the blood, and this ensures the satisfaction of certain conditions (supply of nutriment and removal of waste) which are necessary for the proper working of the organism. (1959, 305)

To formulate the situation more abstractly, Hempel offers the following schema:

(3.4) *Basic pattern of a functional analysis*: The object of the analysis is some "item" i, which is a relatively persistent trait or disposition (e.g., the beating of the heart) occurring in a system s, (e.g., the body of a living vertebrate); and the analysis aims to show that s, is in a state, or internal condition, c_i and in an environment representing certain external conditions c_e such that under conditions c_i and c_e (jointly referred to as c) the trait i has effects which satisfy some "need" or "functional requirement" of s, i.e., a condition n which is necessary for the system's remaining in adequate, or effective, or proper, working order. (1959, 306)

It is clear from this characterization that Hempel, like Nagel, is thinking primarily in terms of the furnace analogy rather than that of the homing torpedo.

Given the fact that Hempel regards explanations as arguments, he seeks an argument that, under the conditions outlined in (3.4), would constitute an explanation of item i (the heartbeat). The crux of the problem that arises is this. We can assert (1) that if i is present and the system is in normal conditions c then n (it will be operating normally). We see (2) that n obtains (it is operating normally in these conditions). If, however, we attempt to deduce the presence of i from premises (1) and (2), we will be guilty of committing the fallacy of affirming the consequent. Given i in the circumstances we can deduce the normal operation n, but given the normal operation n we cannot validly deduce the presence of i. Moreover, Hempel recognizes, there is no inductive argument with (1) and (2) as premises which establishes the presence of i with high inductive probability; thus it cannot qualify as an acceptable I-S explanation either.

The problem to which Hempel points arises quite generally with functional explanations; it is the problem of *functional equivalents*. When we identify some item as fulfilling a function, we recognize that it is sufficient to produce some result in a certain situation. But usually we cannot claim that it is the only possible device that would fulfill that function. It is not necessary for the realization of the goal. As an obvious example, consider the large ears of the jackrabbit. This is one effective mechanism for cooling the animal, but other animals use other mechanisms. Humans perspire and dogs pant. These alternative devices are sufficient to bring about the result in question. Given only that the jackrabbit is an animal that inhabits hot regions, and thus must have some mechanism for reducing body heat, it does not follow deductively, or with high inductive probability, that the jackrabbit has extra-large ears.

Similar remarks apply to the other examples we have considered. Given that a particular ceremony enhances social cohesiveness in a given primitive society, it does not follow that no other ceremony could achieve the same end. Given that a particular dream fulfills a particular unconscious wish, it does not follow that no other dream could have done the job. Given that hemoglobin transports oxygen from one part of the body to another, it does not follow that no other chemical substance could fulfill the same function. Given that the heartbeat causes the blood to circulate, it does not follow that there is no other way for that result to come about. The problem about functional explanation can be stated simply in the following terms. In a correct D-N explanation the explanans is logically sufficient for the explanandum. In the typical functional explanation the explanandum is, given the conditions, sufficient for the explanans. From Hempel's standpoint that is just the wrong way around.

In his discussion of physiological examples, Nagel attempts to avoid the problem of functional equivalents by arguing that, if we specify with sufficient precision the kind of organism we are dealing with, only one kind of mechanism will do the trick. To cite Hempel's example, perhaps there are, in principle, mechanisms other than a beating heart that could circulate blood. Nevertheless, given the stage of evolution of homo sapiens, that is the only mechanism available. Hence if we have a healthy human—not someone undergoing surgery with a heart-lung machine attached—we know that blood is circulating, and if the blood is circulating, we know that the heart is beating. Nagel's prime example is the presence of chlorophyll in green plants (1961, 403–6). To take my example of the jackrabbit's ears, it could be claimed that in the circumstances in which that trait evolved, perspiring and panting were not biologically available. In a hot *dry* climate—the habitat of the jackrabbit— the conservation of water is absolutely critical; both perspiration and panting deplete the animal's supply of water. Whether other heat regulatory mechanisms might be available I do not know.

In view of this treatment of the problem of functional equivalents, Nagel has offered a solution to Hempel's problem about the failure of functional explanations to fit the received models. He can say quite straightforwardly that they do fit the deductive model (Nagel 1961, 405). In 1956 Nagel seemed to claim that Merton's brand of functional analysis would fit the same pattern as functional explanation in physiology, but in 1961 he had doubts that the problem of functional equivalents could be that easily dismissed in the realm of social sciences (Nagel 1961, 533–35). It seems to me, therefore, that Nagel has not succeeded in eliminating the problem of functional equivalents for teleological or functional explanation in general.

Hempel's analysis of functions is, I think, logically impeccable. If an admissible explanation of any fact must be an argument to the effect that the fact-to-be-explained *was to be expected* by virtue of the explanatory facts, then functional 'explanations' are not admissible explanations. But I have often noticed that, in philosophy as well as other human endeavors, one person's counterexample is another's modus ponens. Hempel concludes from his discussion that functional analysis cannot qualify as an admissible type of explanation; at best, it has heuristic value. Others, myself included, would take the moral to be that, since functional explanations play a legitimate scientific role, explanations cannot always be arguments of the sorts endorsed by the received view.[28] I considered it a virtue of the statistical-relevance model that it did not encounter similar problems in connection with functional explanation (W. Salmon 1982).

It turns out that there is a deep connection between the problem of functional or teleological explanation and the problem of explanation in history. Explanations in human history, as well as in the (other?) sciences of human behavior, make frequent appeals to conscious purposes and goals. As Braithwaite pointed out, such cases pose no serious problem for the philosophical theory of scientific explanation.[29] Explanations in these disciplines may, in other cases, make use of unconscious purposes, as in the example of the rain dance. In evolutionary biology functional considerations play a crucial role, and—since the time of Darwin

– it has been appropriate to deny that such appeals to functions involve the conscious purposes of a creator, or any other sort of final causation. The basic idea is that we can understand the evolutionary process in terms of efficient causes in conjunction, perhaps, with chance occurrences. Evolutionary biology thus requires a causal explication of function.

Furthermore, many authors who deal with explanation in human history insist that there is a *narrative* or *genetic* type of explanation that consists in telling the story leading up to the event to be explained.[30] Since the mere recital of just any set of preceding occurrences may have no explanatory value whatever, the narrative must involve events that are causally relevant to the explanandum if it is to serve as an explanation. This, again, demands some form of causal explanation. In human history, as in evolutionary biology, we need an account of what has come to be known as *etiological explanation* that will encompass both genetic and functional explanation.[31] In my opinion, the real breakthrough on functional explanation was provided by Larry Wright (1976) when he advanced an explicitly causal account of functional or teleological explanation.[32] Nagel took the subject up again in 1977. We shall return to this topic in §3.8.

The Second Decade (1958–67)
Manifest Destiny – Expansion and Conflict

As we have already remarked, the first decade after the publication of the Hempel-Oppenheim paper saw little published criticism – or acknowledgment – of it. Quite possibly this portion of the received view – the box in the upper left corner of Table 1 – was accepted with considerable satisfaction for the most part by the philosophy of science community. The situation changed rather dramatically around 1958. This was the year in which the second volume of *Minnesota Studies in the Philosophy of Science* (Feigl et al. 1958) was published, containing Scriven's first article attacking the D-N model. Hanson's *Patterns of Discovery* (1958) appeared during the same year.

In the next few years following 1958 a great many papers on scientific explanation appeared, devoted mainly to issues we have already mentioned. These included, for example, debates on the covering law conception and on the explanation/prediction symmetry thesis. The critiques of the Hempel-Oppenheim account fall into three main categories. First, as we have seen, the Eberle-Kaplan-Montague (1961) critique and replies to it come under the heading of sympathetic efforts to find and eliminate any technical flaws in that explication. Unlike the following two types, these were critiques of the formal explication given in Part III of the Hempel-Oppenheim paper, rather than objections to the preliminary conditions of adequacy advanced in Part I. Second, as we shall see in §2.1, the attacks by Hanson, Scriven, and others were motivated by deep philosophical disagreements with anything resembling the logical empiricist point of view. Third, there were constructive efforts by philosophers such as Bromberger (1962, 1963, 1966) and Scheffler (1957, 1963) who generally accepted something like the received view, and sought ways to improve and perfect it. As we shall see in §2.3, many of the problems raised under the latter two headings were formulated with the aid of putative counterexamples that have since become standard in the literature.

During this time there were also attempts to further elaborate or defend the received view. One major effort in that direction can be found in May Brodbeck's (1962) defense of deductive explanation in her contribution to the third volume