

## *Four Decades of Scientific Explanation*

### *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.<sup>1</sup> 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.<sup>2</sup>

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.<sup>3</sup>

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.<sup>4</sup>

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.<sup>5</sup> 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 statisti-

cal laws, such 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	Particular Facts	General Regularities
Laws		
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.<sup>6</sup> 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.