INTRODUCTION TO THE PHILOSOPHY OF SCIENCE

A Text by Members of the Department of the History and Philosophy of Science of the University of Pittsburgh

Merrilee H. Salmon · John Earman · Clark Glymour
James G. Lennox · Peter Machamer · J. E. McGuire
John D. Norton · Wesley C. Salmon · Kenneth F. Schaffner

Hackett Publishing Company
Indianapolis/Cambridge

To the giants
On whose shoulders we stand

Reprinted 1999 by Hackett Publishing Company, Inc.

Printed in the United States of America

05 04 03 02 01 00 99   1 2 3 4 5 6 7 8 9

For further information, please address
Hackett Publishing Company, Inc.
PO Box 44937
Indianapolis, IN 46244-0937

Library of Congress Cataloging-in-Publication Data

Introduction to the philosophy of science / Merrilee H. Salmon . . . [et al.].
p. cm.
Originally published: Englewood Cliffs, N.J. : Prentice Hall, c 1992.
Includes bibliographical references and index.
ISBN 0-87220-451-0 (cloth). — ISBN 0-87220-450-2 (paper)
1. Science—Philosophy. I. Salmon, Merrilee H.
Q175.I633 1999   98-48586
501—dc21   CIP

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

One

Scientific Explanation

Wesley C. Salmon

The fruits of science are many and various. When science is first mentioned, many people think immediately of high technology. Such items as computers, nuclear energy, genetic engineering, and high-temperature superconductors are likely to be included. These are the fruits of applied science. Evaluating the benefits, hazards, and costs of such technological developments often leads to lively and impassioned debate.

In this chapter, however, we are going to focus on a different aspect of science, namely, the intellectual understanding it gives us of the world we live in. This is the fruit of pure science, and it is one we highly prize. All of us frequently ask the question "Why?" in order to achieve some degree of understanding with regard to various phenomena. This seems to be an expression of a natural human curiosity. Why, during a total lunar eclipse, does the moon take on a coppery color instead of just becoming dark when the earth passes between it and the sun? Because the earth's atmosphere acts like a prism, diffracting the sunlight passing through it in such a way that the light in the red region of the spectrum falls upon the lunar surface. This is a rough sketch of a scientific explanation of that phenomenon, and it imparts at least some degree of scientific understanding.

Our task in this chapter is to try to say with some precision just what scientific explanation consists in. Before we embark on that enterprise, however, some preliminary points of clarification are in order.

1.1 EXPLANATION VS. CONFIRMATION

The first step in clarifying the notion of scientific explanation is to draw a sharp distinction between explaining *why* a particular phenomenon occurs and giving reasons for believing *that* it occurs. My reason for believing *that* the moon turns coppery during total eclipse is that I have observed it with my own eyes. I can also appeal to the testimony of other observers. That is how the proposition that the moon turns coppery during a total eclipse is confirmed,[1] and it is entirely different from explaining *why* it happens. Consider another example. According to contemporary cosmology all of the distant galaxies are receding from us at high velocities. The evidence for this is the fact that the light from them is shifted toward the red end of the spectrum; such evidence confirms the statement that the other galaxies are moving away from our galaxy (the Milky Way). The fact that there is such a red shift does not explain *why* the galaxies are moving in that way; instead, the fact that they are receding explains—in terms of the Doppler effect[2]—why the light is shifted toward the red end of the spectrum. The explanation of the recession lies in the "big bang" with which our universe began several billion years ago; this is what makes all of the galaxies recede from one another and, consequently, makes all of the others move away from us.

1.2 OTHER KINDS OF EXPLANATION

Another preliminary step in clarifying the notion of *scientific* explanation is to recognize that there are many different kinds of explanation in addition to those we classify as scientific. For example, we often encounter explanations of *how* to do something—how to use a new kitchen gadget, or how to find a certain address in a strange city. There are, in addition, explanations of *what*—what an unfamiliar word means, or what is wrong with an automobile. While many, if not all, scientific explanations can be requested by means of why-questions, requests for explanations of these other sorts would not normally be phrased in why-questions; instead, *how-to*-questions and *what*-questions would be natural.

Still other types of explanation exist. Someone might ask for an explanation of the meaning of a painting or a poem; such a request calls for an artistic interpretation. Or, someone might ask for an explanation of a mathematical proof; an appropriate response would be to fill in additional steps to show how one gets from one step to another in the original demonstration. Neither of these qualifies as scientific explanation. Also excluded from our domain of scientific explanation are explanations of formal facts of pure mathematics, such as the infinitude of the set of prime numbers. We are concerned only with explanation in the empirical sciences.

As we understand the concept of *scientific* explanation, such an explanation is an attempt to render understandable or intelligible some particular event (such as the 1986 accident at the Chernobyl nuclear facility) or some general fact (such as the copper color of the moon during total eclipse) by appealing to other particular and/or general facts drawn from one or more branches of empirical science. This formulation

[1] Confirmation will be treated in Chapter 2 of this book.

[2] The Doppler effect is the lengthening of waves emitted by a source traveling away from a receiver and the shortening of waves emitted by a source approaching a receiver. This effect occurs in both light and sound, and it can be noticed in the change of pitch of a whistle of a passing train.

is *not* meant as a definition, because such terms as "understandable" and "intelligible" are as much in need of clarification as is the term "explanation." But it should serve as a rough indication of what we are driving at.

In pointing out the distinction between scientific explanations and explanations of other types we do not mean to disparage the others. The aim is only to emphasize the fact that the word "explanation" is extremely broad—it applies to a great many different things. We simply want to be clear on the type of explanation with which our discussion is concerned.

1.3 SCIENTIFIC EXPLANATIONS AND WHY-QUESTIONS

Many scientific explanations are requested by means of why-questions, and even when the request is not actually formulated in that way, it can often be translated into a why-question. For example, "What caused the Chernobyl accident?" or "For what reason did the Chernobyl accident occur?" are equivalent to "Why did the Chernobyl accident occur?" However, not all why-questions are requests for scientific explanations. A woman employee might ask why she received a smaller raise in salary than a male colleague when her job-performance is just as good as his. Such a why-question might be construed as a request for a justification, or, perhaps, simply a request for more pay. A bereaved widow might ask why her husband died even though she fully understands the medical explanation. Such a why-question is a request for consolation, not explanation. Some why-questions are requests for evidence. To the question, "Why should we believe that the distant galaxies are traveling away from us at high velocities?" the answer, briefly, is the red shift. Recall, as we noted in Section 1.1, that the red shift does not explain the recession. The recession explains the red shift; the red shift is evidence for the recession. For the sake of clarity we distinguish *explanation-seeking why-questions* from why-questions that seek such other things as justification, consolation, or evidence.

Can all types of scientific explanation be requested by why-questions? Some authors say "yes" and others say "no." It has been suggested, for example, that some scientific explanations are answers to *how-possibly-questions*. There is an old saying that a cat will always land on its feet (paws), no matter what position it falls from. But remembering the law of conservation of angular momentum we might well ask, "How is it possible for a cat, released (without imparting any angular momentum) from a height of several feet above the ground with its legs pointing upward, to turn over so that its paws are beneath it when it lands? Is this just an unfounded belief with no basis in fact?" The answer is that the cat can (and does) twist its body in ways that enable it to turn over without ever having a total angular momentum other than zero (see Frohlich 1980).

Other requests for explanation may take a *how-actually* form. A simple commonsense example illustrates the point. "How did the prisoner escape?" calls for an explanation of how he did it, not why he did it. The answer to this question might be that he sawed through some metal bars with a hacksaw blade smuggled in by his wife. If we were to ask why, the answer might be his intense desire to be with his wife outside of the prison. For a somewhat more scientific example, consider the question, ''How did large mammals get to New Zealand?'' The answer is that they came in boats—the first were humans, and humans brought other large mammals. Or, consider the question, ''How is genetic information transmitted from parents to offspring?'' The answer to this question involves the structure of the DNA molecule and the genetic code.

In this chapter we will not try to argue one way or other on the issue of whether all scientific explanations can appropriately be requested by means of why-questions. We will leave open the possibility that some explanations cannot suitably be requested by why-questions.

1.4 SOME MATTERS OF TERMINOLOGY

As a further step in preliminary clarification we must establish some matters of terminology. In the first place, any explanation consists of two parts, the *explanandum* and the *explanans*. The explanandum is the fact that is to be explained. This fact may be a *particular* fact, such as the explosion of the Challenger space-shuttle vehicle, or a *general* fact, such as the law of conservation of linear momentum. A statement to the effect that the explanandum obtained is called the *explanandum-statement*. Sometimes, when it is important to contrast the fact-to-be-explained with the statement of the explanandum, we may refer to the explanandum itself as the *explanandum-fact*. When the explanandum is a particular fact we often speak of it as an event or occurrence, and there is no harm in this terminology, provided we are clear on one basic point. By and large, the events that happen in our world are highly complex, and we hardly ever try to explain every aspect of such an occurrence. For example, in explaining the explosion of the Challenger vehicle, we are not concerned to explain the fact that a woman was aboard, the fact that she was a teacher, or the fact that her life had been insured for a million dollars. When we speak of a particular fact, it is to be understood that this term refers to certain limited aspects of the event in question, not to the event in its full richness and complexity.

The other part of an explanation is the explanans. The explanans is that which does the explaining. It consists of whatever facts, particular or general, are summoned to explain the explanandum. When we want to refer to the statements of these facts we may speak of the *explanans-statements*; to contrast the facts with the statements of them we may also speak of the *explanans-facts*.

In the philosophical literature on scientific explanation, the term ''explanation'' is used ambiguously. Most authors use it to apply to a linguistic entity composed of the explanans-statements and the explanandum-statement. Others use it to refer to the collection of facts consisting of the explanans-facts and the explanandum-fact. In most contexts this ambiguity is harmless and does not lead to any confusion. But we should be aware that it exists.

1.5 DEDUCTION AND INDUCTION

As we will see, one influential philosophical account of explanation regards all bona fide scientific explanations as arguments. An argument is simply a set of statements, one of which is singled out as the conclusion of the argument. The remaining members of the set are premises. There may be one or more premises; no fixed number of premises is required.[3] The premises provide support for the conclusion.

All logically correct arguments fall into two types, deductive and inductive, and these types differ fundamentally from one another. For purposes of this chapter (and later chapters as well) we need a reasonably precise characterization of them. Four characteristics are important for our discussion.

DEDUCTION	INDUCTION
1. In a valid deductive argument, all of the content of the conclusion is present, at least implicitly, in the premises. Deduction is *nonampliative.*	1. Induction is *ampliative.* The conclusion of an inductive argument has content that goes beyond the content of its premises.
2. If the premises are true, the conclusion must be true. Valid deduction is *necessarily truth-preserving.*	2. A correct inductive argument may have true premises and a false conclusion. Induction is *not necessarily truth-preserving.*
3. If new premises are added to a valid deductive argument (and none of the original premises is changed or deleted) the argument remains valid. Deduction is *erosion-proof.*	3. New premises may completely undermine a strong inductive argument. Induction is *not erosion-proof.*
4. Deductive validity is an *all-or-nothing* matter; validity does not come in degrees. An argument is totally valid or it is invalid.	4. Inductive arguments come in different *degrees of strength.* In some inductions the premises support the conclusions more strongly than in others.

These characteristics can be illustrated by means of simple time-honored examples.

(1) All humans are mortal.
Socrates is human.
―――――――――――
Socrates is mortal.

Argument (1) is obviously a valid deduction. When we have said that all humans are mortal, we have already said that Socrates is mortal, given that Socrates is human. Thus, it is nonampliative. Because it is nonampliative, it is necessarily truth-preserving. Since nothing is said by the conclusion that is not already stated by the premises, what the conclusion says *must* be true if what the premises assert is true. Moreover, the argument remains nonampliative, and hence, necessarily truth-preserving, if new premises—for example, "Xantippe is human"—are added. You cannot make a valid deduction invalid just by adding premises. Finally, the premises

[3] Because of certain logical technicalities, there are valid deductive arguments that have no premises at all, but arguments of this sort will not be involved in our discussion.

support the conclusion totally, not just to some degree; to accept the premises and reject the conclusion would be outright self-contradiction.

(2) All observed ravens have been black.
$\overline{\overline{}}$
All ravens are black.

This argument is obviously ampliative; the premise refers only to ravens that have been observed, while the conclusion makes a statement about all ravens, observed or unobserved. It is not necessarily truth-preserving. Quite possibly there is, was, or will be—at some place or time—a white raven, or one of a different color. It is not erosion-proof; the observation of one non-black raven would undermine it completely. And its strength is a matter of degree. If only a few ravens in one limited environment had been observed, the premise would not support the conclusion very strongly; if vast numbers of ravens have been observed under a wide variety of circumstances, the support would be much stronger. But in neither case would the conclusion be necessitated by the premise.

Deductive validity and inductive correctness do not hinge on the truth of the premises or the conclusion of the argument. A valid deduction may have true premises and a true conclusion, one or more false premises and a false conclusion, and one or more false premises and a true conclusion.[4] When we say that valid deduction is necessarily truth-preserving, we mean that the conclusion would have to be true *if the premises were true*. Thus there cannot be a valid deduction with true premises and a false conclusion. Where correct inductive arguments are concerned, since they are not necessarily truth-preserving, any combination of truth values of premises and conclusion is possible. What we would like to say is that, if the premises are true (and embody all relevant knowledge), the conclusion is probable. As we will see in Chapter 2, however, many profound difficulties arise in attempting to support this claim about inductive arguments.

We have chosen very simple—indeed, apparently trivial—examples in order to illustrate the basic concepts. In actual science, of course, the arguments are much more complex. Most of the deductive arguments found in serious scientific contexts are mathematical derivations, and these can be extremely complicated. Nevertheless, the basic fact remains that all of them fulfill the four characteristics listed above. Although deep and interesting problems arise in the philosophy of mathematics, they are not our primary concern in this book. Our attention is focused on the empirical sciences, which, as we argue in Chapter 2, necessarily involve induction. In that chapter we encounter much more complex and interesting inductive arguments.

1.6 IS THERE ANY SUCH THING AS SCIENTIFIC EXPLANATION?

The idea that science can furnish explanations of various phenomena goes back to Aristotle (4th century B.C.), and it has been reaffirmed by many philosophers and

[4] The familiar slogan, "Garbage in, garbage out," does not accurately characterize deductive arguments.

scientists since then. Nevertheless, many other philosophers and scientists have maintained that science must "stick to the facts," and consequently can answer only questions about *what* but not about *why*. To understand "the *why* of things," they felt, it is necessary to appeal to theology or metaphysics. Science can describe natural phenomena and predict future occurrences, but it cannot furnish explanations. This attitude was particularly prevalent in the early decades of the twentieth century. Since it is based upon certain misconceptions regarding scientific explanation, we need to say a bit about it.

It is natural enough, when attempting to find out why a person did something, to seek a conscious (or perhaps unconscious) motive. For example, to the question, "Why did you buy that book?" a satisfactory answer might run, "Because I wanted to read an amusing novel, and I have read several other novels by the same author, all of which I found amusing." This type of explanation is satisfying because we can put ourselves in the place of the subject and *understand* how such motivation works. The concept of *understanding* is critical in this context, for it signifies empathy. If we yearn for that kind of empathetic understanding of nonhuman phenomena, we have to look elsewhere for motivation or purpose. One immediate suggestion is to make the source of purpose supernatural. Thus, prior to Darwin, the variety of species of living things was explained by *special creation*—that is, God's will. Another manifestation of the same viewpoint—held by some, but not all, *vitalists*—was the notion that behind all living phenomena there is a vital force or *entelechy* directing what goes on. These entities—entelechies and vital forces—are not open to empirical investigation.

The insistence that all aspects of nature be explained in human terms is known as *anthropomorphism*. The supposition—common before the rise of modern science—that the universe is a cozy little place, created for our benefit, with humans at its center, is an anthropomorphic conception. The doctrines of special creation and some forms of vitalism are anthropomorphic. So-called "creation science" is anthropomorphic. Teleological explanation of nonhuman phenomena in terms of human-like purposes is anthropomorphic.[5]

Many philosophers and scientists rejected the appeal to anthropomorphic and teleological explanations as an appeal to hypotheses that could not, even in principle, be investigated by empirical science. If this is what is needed for explanation, they said, we want no part of it. Science is simply not concerned with explaining natural phenomena; anyone who wants explanations will have to look for them outside of science. Such scientists and philosophers were eager to make clear that scientific knowledge does not rest on nonempirical metaphysical principles.

Not all philosophers were willing to forgo the claim that science provides explanations of natural phenomena. Karl R. Popper (1935), Carl G. Hempel (1948), R. B. Braithwaite (1953), and Ernest Nagel (1961) published important works in which they maintained that there are such things as legitimate scientific explanations, and that such explanations can be provided without going beyond the bounds of empirical science. They attempted to provide precise characterizations of scientific

[5] As James Lennox points out in Chapter 7, teleological explanations are anthropomorphic *only* if they appeal to human-like purposes. In evolutionary biology—and other scientific domains as well—there are teleological explanations that are *not* anthropomorphic.

explanation, and they were, to a very large degree, in agreement with respect to the core of the account. The line of thought they pursued grew into a theory that enjoyed a great deal of acceptance among philosophers of science. We will discuss it at length in later sections of this chapter.

1.7 DOES EXPLANATION INVOLVE REDUCTION TO THE FAMILIAR?

It has sometimes been asserted that explanation consists in reducing the mysterious or unfamiliar to that which is familiar. Before Newton, for example, comets were regarded as mysterious and fearsome objects. Even among educated people, the appearance of a comet signified impending disaster, for example, earthquakes, floods, famines, or epidemic diseases. Newton showed that comets could be understood as planet-like objects that travel around the sun in highly eccentric orbits. For that reason, any given comet spends most of its time far from the sun and well beyond the range of human observation. When one appeared it was a surprise. But when we learned that they behave very much as the familiar planets do, their behavior was explained, and they were no longer objects of dread.

Appealing as the notion of reduction of the unfamiliar to the familiar may be, it is not a satisfactory characterization of scientific explanation. The point can best be made in terms of a famous puzzle known as *Olbers's paradox*—which is named after a nineteenth-century astronomer but was actually formulated by Edmund Halley in 1720—why is the sky dark at night? Nothing could be more familiar than the darkness of the night sky. But Halley and later astronomers realized that, if Newton's conception of the universe were correct, then the whole night sky should shine as brightly as the noonday sun. The question of how to explain the darkness of the sky at night is extremely difficult, and there may be no answer generally accepted by the experts. Among the serious explanations that have been offered, however, appeal is made to such esoteric facts as the non-Euclidean character of space or the mean free path of photons in space. In this case, and in many others as well, a familiar phenomenon is explained by reference to facts that are very unfamiliar indeed.

I suspect that a deep connection exists between the anthropomorphic conception of explanation and the thesis that explanation consists in reduction of the unfamiliar to the familiar. The type of explanation with which we are best acquainted is that in which human action is explained in terms of conscious purposes. If it is possible to explain the phenomena of physics or biology in terms of attempting to realize a goal, that is a striking case of reduction to the familiar. A problem with this approach is, of course, that a great deal of the progress in scientific understanding has resulted in the elimination, not the injection, of purposes.

1.8 THE DEDUCTIVE-NOMOLOGICAL PATTERN OF SCIENTIFIC EXPLANATION

In a classic 1948 paper, Carl G. Hempel and Paul Oppenheim formulated, with great precision, one pattern of scientific explanation that is central to all discussions of the subject. It is known as the *deductive-nomological* (*D-N*) *model* of scientific explanation. Stated very simply, an explanation of this type explains by subsuming its explanandum-fact under a general law. This can best be appreciated by looking at an example.

A figure skater with arms outstretched stands balanced on one skate. Propelling herself with her other skate she begins to rotate slowly. She stops propelling herself, but she continues to rotate slowly for a few moments. Suddenly—without propelling herself again and without being propelled by any external object, such as another skater—she begins spinning very rapidly. Why? Because she drew her arms in close to her body, thus concentrating her total body mass closer to the axis of rotation. Because of the law of conservation of angular momentum, her rate of rotation had to increase to compensate for her more compact body configuration.

More technically, the angular momentum of an object is the product of its angular velocity (rate of rotation) and its moment of inertia. The moment of inertia depends upon the mass of the object and the average distance of the mass from the axis of rotation; for a fixed mass, the moment of inertia is smaller the more compactly the mass is distributed about the axis of rotation. The law of conservation of angular momentum says that the angular momentum of a body that is not being propelled or retarded by external forces does not change; hence, since the moment of inertia is decreased, the rate of rotation must increase to keep the value of the product constant.[6]

According to Hempel and Oppenheim, an explanation of the foregoing sort is to be viewed as a deductive argument. It can be set out more formally as follows:

(3) The angular momentum of any body (whose rate of rotation is not being increased or decreased by external forces) remains constant.
The skater is not interacting with any external object in such a way as to alter her angular velocity.
The skater is rotating (her angular momentum is not zero).
The skater reduces her moment of inertia by drawing her arms in close to her body.

The skater's rate of rotation increases.

The explanandum—the increase in the skater's rate of rotation—is the conclusion of the argument. The premises of the argument constitute the explanans. The first premise states a law of nature—the law of conservation of angular momentum. The remaining three premises state the antecedent conditions. The argument is logically correct; the conclusion follows validly from the premises. For purposes of our discussion, we may take the statements of antecedent conditions as true; expert figure skaters do this maneuver frequently. The law of conservation of angular momentum can also be regarded as true since it is a fundamental law of physics which has been confirmed by a vast quantity of empirical data.

Hempel and Oppenheim set forth four conditions of adequacy for D-N explanations:

[6] In this example we may ignore the friction of the skate on the ice, and the friction of the skater's body in the surrounding air.

1. The explanandum must be a logical consequence of the explanans; that is, the explanation must be a valid deductive argument.
2. The explanans must contain at least one general law, and it must actually be required for the derivation of the explanandum; in other words, if the law or laws were deleted, without adding any new premises, the argument would no longer be valid.
3. The explanans must have empirical content; it must be capable, at least in principle, of test by experiment or observation.
4. The sentences constituting the explanans must be true.

These conditions are evidently fulfilled by our example. The first three are classified as *logical* conditions of adequacy; the fourth is *empirical*. An argument that fulfills all four conditions is an explanation (for emphasis we sometimes say "true explanation"). An argument that fulfills the first three conditions, without necessarily fulfilling the fourth, is called a *potential explanation*. It is an argument that would be an explanation if its premises were true.[7]

According to Hempel and Oppenheim, it is possible to have D-N explanations, not only of particular occurrences as in argument (3), but also of general laws. For example, in the context of Newtonian mechanics, it is possible to set up the following argument:

(4) $F = ma$ (Newton's second law).
For every action there is an equal and opposite reaction (Newton's third law).

In every interaction, the total linear momentum of the system of interacting bodies remains constant (law of conservation of linear momentum).

This argument is valid, and among its premises are statements of general laws. There are no statements of antecedent conditions, but that is not a problem since the conditions of adequacy do not require them. Because we are not concerned to explain any particular facts, no premises regarding particular facts are needed. Both premises in the explanans are obviously testable, for they have been tested countless times. Thus, argument (4) fulfills the logical conditions of adequacy, and consequently, it qualifies as a potential explanation. Strictly speaking, it does not qualify as a true explanation, for we do not consider Newton's laws of motion literally true, but in many contexts they can be taken as correct because they provide extremely accurate approximations to the truth.

Although Hempel and Oppenheim discussed both deductive explanations of particular facts and deductive explanations of general laws, they offered a precise characterization only of the former, but not of the latter. They declined to attempt to provide a characterization of explanations of general laws because of a problem they recognized but did not know how to solve. Consider Kepler's laws of planetary motion K and Boyle's law of gases B. If, on the one hand, we conjoin the two to form

[7] Hempel and Oppenheim provide, in addition to these conditions of adequacy, a precise technical definition of "explanation." In this book we will not deal with these technicalities.

a law $K \,.\, B$, we can obviously deduce K from it. But this could not be regarded as an explanation of K, for it is only a pointless derivation of K from itself. On the other hand, the derivation of K from Newton's laws of motion and gravitation constitutes an extremely illuminating explanation of Kepler's laws. Hempel and Oppenheim themselves confessed that they were unable to provide any criterion to distinguish the pointless pseudoexplanations from the genuine explanations of laws (see Hempel and Oppenheim 1948 as reprinted in Hempel 1965b, 273, f.n. 33).

Hempel and Oppenheim envisioned two types of D-N explanation, though they were able to provide an account of only one of them. In addition, they remarked that other types of explanation are to be found in the sciences, namely, explanations that appeal, not to universal generalizations, but to statistical laws instead (ibid., 250–251).

Table 1.1 shows the four kinds of explanations to which Hempel and Oppenheim called attention; they furnished an account only for the type found in the upper left-hand box. Some years later Hempel (1962) offered an account of the I-S pattern in the lower left-hand box. In Hempel (1965b) he treated both the I-S and the D-S patterns. In 1948, Hempel and Oppenheim were looking forward to the time when theories of explanation dealing with all four boxes would be available.

TABLE 1.1

Explanada / 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

1.9 WHAT ARE LAWS OF NATURE?

Hempel and Oppenheim emphasized the crucial role played by laws in scientific explanation; in fact, the D-N pattern is often called *the covering-law model*. As we will see, laws play a central part in other conceptions of scientific explanation as well. Roughly speaking, a *law* is a regularity that holds throughout the universe, at all places and all times. A *law-statement* is simply a statement to the effect that such a regularity exists. A problem arises immediately. Some regularities appear to be lawful and others do not. Consider some examples of laws:

(i) All gases, kept in closed containers of fixed size, exert greater pressure when heated.
(ii) In all closed systems the quantity of energy remains constant.
(iii) No signals travel faster than light.

Contrast these with the following:

(iv) All of the apples in my refrigerator are yellow.
(v) All Apache basketry is made by women.
(vi) No golden spheres have masses greater than 100,000 kilograms.

Let us assume, for the sake of argument, that all of the statements (i)–(vi) are true. The first thing to notice about them is their generality. Each of them has the overall form, "All *A* are *B*" or "No *A* are *B*." Statements having these forms are known as *universal generalizations*. They mean, respectively, "*Anything* that is an *A* is also a *B*" and "*Nothing* that is an *A* is also a *B*." Nevertheless, statements (i)–(iii) differ fundamentally from (iv)–(vi). Notice, for example, that none of the statements (i)–(iii) makes any reference to any particular object, event, person, time, or place. In contrast, statement (iv) refers to a particular person (me), a particular object (my refrigerator), and a particular time (now). This statement is not completely general since it singles out certain particular entities to which it refers. The same remark applies to statement (v) since it refers to a particular limited group of people (the Apache).

Laws of nature are generally taken to have two basic capabilities. First, they support counterfactual inferences. A *counterfactual statement* is a conditional statement whose antecedent is false. Suppose, for example, that I cut a branch from a tree and then, immediately, burn it in my fireplace. This piece of wood was never placed in water and never will be. Nevertheless, we are prepared to say, without hesitation, that *if it had been placed in water, it would have floated*. This italicized statement is a counterfactual conditional. Now, a law-statement, such as (i), will support a counterfactual assertion. We can say, regarding a particular sample of some gas, held in a closed container of fixed size but not actually being heated, that *if* it were heated it *would* exert greater pressure. We can assert the counterfactual because we take statement (i) to be a statement of a law of nature.

When we look at statement (iv) we see that it does not support any such counterfactual statement. Holding a red delicious apple in my hand, I cannot claim, on the basis of (iv), that this apple would be yellow if it were in my refrigerator.

A second capability of laws of nature is to support *modal* statements of physical necessity and impossibility. Statement (ii), the first law of thermodynamics, implies that it is impossible to create a perpetual motion machine of the first kind—that is, a machine that does useful work without any input of energy from an external source. In contrast, statement (v) does not support the claim that it is impossible for an Apache basket to be made by a male. It is physically possible that an Apache boy might be taught the art of basket making, and might grow up to make a career of basketry.

When we compare statements (iii) and (vi) more subtle difficulties arise. Unlike statements (iv) and (v), statement (vi) does not make reference to any particular entity or place or time.[8] It seems clear, nevertheless, that statement (vi)—even assuming it to be true—cannot support either modal statements or counterfactual

[8] If the occurrence of the kilogram in (vi) seems to make reference to a particular object—the international prototype kilogram kept at the international bureau of standards—the problem can easily be circumvented by defining mass in terms of atomic mass units.

conditionals. Even if we agree that nowhere in the entire history of the universe—past, present, or future—does there exist a gold sphere of mass greater than 100,000 kilograms, we would not be justified in claiming that it is *impossible* to fabricate a gold sphere of such mass. I once made a rough calculation of the amount of gold in the oceans of the earth, and it came to about 1,000,000 kilograms. If an incredibly rich prince were determined to impress a woman passionately devoted to golden spheres it would be physically possible for him to extract a little more than 100,000 kilograms from the sea to create a sphere that massive.

Statement (vi) also lacks the capacity to support counterfactual conditionals. We would not be justified in concluding that, if two golden hemispheres, each of 50,001 kilogram mass, were put together, they would not form a golden sphere of mass greater than 100,000 kilograms. To appreciate the force of this point, consider the following statement:

(vii) No enriched uranium sphere has a mass greater than 100,000 kilograms.

This *is* a lawful generalization, because the critical mass for a nuclear chain reaction is just a few kilograms. If 100,000 kilograms of enriched uranium were to be assembled, we would have a gigantic nuclear explosion. No comparable catastrophe would ensue, as far as we know, if a golden sphere of the same mass were put together.

Philosophers have often claimed that we can distinguish true generalizations that are *lawful* from those that are *accidental*. Even if we grant the truth of (vi), we must conclude that it is an accidental generalization. Moreover, they have maintained that among universal generalizations, regardless of truth, it is possible to distinguish *lawlike generalizations* from those that are not lawlike. A lawlike generalization is one that has all of the qualifications for being a law except, perhaps, being true.

It is relatively easy to point to the characteristic of statements (iv) and (v) that makes them nonlawlike, namely, that they make reference to particular objects, persons, events, places, or times. The nonlawlike character of statement (vi) is harder to diagnose. One obvious suggestion is to apply the criteria of supporting counterfactual and/or modal statements. We have seen that (vi) fails on that score. The problem with that approach is that it runs a serious risk of turning out to be circular. Consider statement (ii). Why do we consider it *physically impossible* to build a perpetual motion machine (of the first type)? Because to do so would violate a law of nature, namely (ii). Consider statement (vi). Why do we consider it *physically possible* to fabricate a golden sphere whose mass exceeds 100,000 kilograms? Because to do so would not violate a law of nature. It appears that the question of what modal statements to accept hinges on the question of what regularities qualify as laws of nature.

A similar point applies to the support of counterfactual conditionals. Consider statement (i). Given a container of gas that is not being heated, we can say that, if it were to be heated, it would exert increased pressure on the walls of its container—sufficient in many cases to burst the container. (I learned my lesson on this as a Boy Scout heating an unopened can of beans in a camp fire.) The reason that we can make such a counterfactual claim is that we can infer from statement (i) what would happen, and (i) states a law of nature. Similarly, from (iii) we can deduce that if something travels faster than light it is not a signal—that is, it cannot transmit information. You might think that this is vacuous because, as the theory of relativity tells us, nothing can travel faster than light. However, this opinion is incorrect. Shadows and various other kinds of "things" can easily be shown to travel faster than light. We can legitimately conclude that, if something does travel faster than light, it is not functioning as a signal, because (iii) is, indeed, a law of nature.

What are the fundamental differences between statement (vi) on the one hand and statements (i)–(iii) and (vii) on the other? The main difference seems to be that (i)–(iii) and (vii) are all deeply embedded in well-developed scientific theories, and that they have been, directly or indirectly, extensively tested. This means that (i)–(iii) and (vii) have a very different status within our body of scientific knowledge than do (iv)–(vi). The question remains, however, whether the regularities described by (i)–(iii) and (vii) have a different status in the physical universe than do (iv)–(vi).

At the very beginning of this chapter, we considered the explanation of the fact that the moon assumes a coppery hue during total eclipse. This is a regularity found in nature, but is it a lawful regularity? Is the statement, "The moon turns a coppery color during total eclipses," a law-statement? The immediate temptation is to respond in the negative, for the statement makes an explicit reference to a particular object, namely, our moon. But if we reject that statement as a lawful generalization, it would seem necessary to reject Kepler's laws of planetary motion as well, for they make explicit reference to our solar system. Galileo's law of falling bodies would also have to go, for it refers to things falling near the surface of the earth. It would be unreasonable to disqualify all of them as laws.

We can, instead, make a distinction between basic and derived laws. Kepler's laws and Galileo's law can be derived from Newton's laws of motion and gravitation, in conjunction with descriptions of the solar system and the bodies that make it up. Newton's laws are completely general and make no reference to any particular person, object, event, place, or time. The statement about the color of the moon during total eclipse can be derived from the laws of optics in conjunction with a description of the earth's atmosphere and the configuration of the sun, moon, and earth when an eclipse occurs. The statement about the color of the moon can also be taken as a derivative law. But what about statements (iv) and (v)? The color of the apples in my refrigerator can in no way be derived from basic laws of nature in conjunction with a description of the refrigerator. No matter how fond I may be of golden delicious apples, there is no physical impossibility of a red delicious getting into my refrigerator. Similarly, there are no laws of nature from which, in conjunction with descriptions of the Apache and their baskets, it would be possible to derive that they can only be made by women.

1.10 PROBLEMS FOR THE D-N PATTERN OF EXPLANATION

Quite remarkably the classic article by Hempel and Oppenheim received virtually no attention for a full decade. Around 1958, however, a barrage of criticism began and a lively controversy ensued. Much of the criticism was brought into sharp focus by means of counterexamples that have, themselves, become classic. These examples fall into two broad categories. The first consists of arguments that fulfill all of the requirements for D-N explanation, yet patently fail to qualify as bona fide explanations. They show that the requirements set forth by Hempel and Oppenheim are not *sufficient* to determine what constitutes an acceptable scientific explanation. The second consists of examples of allegedly bona fide explanations that fail to fulfill the Hempel-Oppenheim requirements. They are meant to show that it is not *necessary* to fulfill those requirements in order to have correct explanations. We must treat this second category with care, for Hempel and Oppenheim never asserted that all correct explanations fit the D-N pattern. They explicitly acknowledged that legitimate statistical explanations can be found in science. So, statistical explanations are not appropriate as counterexamples. However, the attempt has been to find examples that are clearly not statistical, but which fail to fulfill the Hempel-Oppenheim criteria. Let us look at some counterexamples of each type.

CE-1. The flagpole and its shadow.[9] On a flat and level piece of ground stands a flagpole that is 12′ tall. The sun, which is at an elevation of 53.13° in the sky, shines brightly. The flagpole casts a shadow that is 9′ long. If we ask why the shadow has that length, it is easy to answer. From the elevation of the sun, the height of the flagpole, and the rectilinear propagation of light, we can deduce, with the aid of a bit of trigonometry, the length of the shadow. The result is a D-N explanation that most of us would accept as correct. So far, there is no problem.

If, however, someone asks why the flagpole is 12′ tall, we could construct essentially the same argument as before. But instead of deducing the length of the shadow from the height of the flagpole and the elevation of the sun, we would deduce the height of the flagpole from the length of the shadow and the elevation of the sun. Hardly anyone would regard that argument, which satisfies all of the requirements for a D-N explanation, as an adequate explanation of the height of the flagpole.

We can go one step farther. From the length of the shadow and the height of the flagpole, using a similar argument, we can deduce that the sun is at an elevation of 53.13°. It seems most unreasonable to say that the sun is that high in the sky because a 12′ flagpole casts a 9′ shadow. From the fact that a 12′ flagpole casts a 9′ shadow we can infer *that* the sun is that high in the sky, but we cannot use those data to explain *why* it is at that elevation. Here we must be sure to remember the distinction between confirmation and explanation (discussed in Section 1.1). The explanation of the elevation rests upon the season of the year and the time of day.

The moral: The reason it is legitimate to explain the length of the shadow in terms of the height of the flagpole and the elevation of the sun is that the shadow is the effect of those two causal factors. We can explain effects by citing their causes. The reason it is illegitimate to explain the height of the flagpole by the length of the shadow is that the length of the shadow is an effect of the height of the flagpole (given the elevation of the sun), but it is no part of the cause of the height of the flagpole. We cannot explain causes in terms of their effects. Furthermore, although the eleva-

[9] The counterexample was devised by Sylvain Bromberger, but to the best of my knowledge he never published it.

tion of the sun is a crucial causal factor in the relation between the height of the flagpole and the length of the shadow, the flagpole and its shadow play no causal role in the position of the sun in the sky.

CE-2. The barometer and the storm. Given a sharp drop in the reading on a properly functioning barometer, we can predict that a storm will shortly occur. Nevertheless, the reading on the barometer does not explain the storm. A sharp drop in atmospheric pressure, which is registered on the barometer, explains both the storm and the barometric reading.

The moral: Many times we find two effects of a common cause that are correlated with one another. In such cases we do not explain one effect by means of the other. The point is illustrated also by diseases. A given illness may have many different symptoms. The disease explains the symptoms; one symptom does not explain another.

CE-3. A solar eclipse. From the present positions of the earth, moon, and sun, using laws of celestial mechanics, astronomers can predict a future total eclipse of the sun. After the eclipse has occurred, the very same data, laws, and calculations provide a legitimate D-N explanation of the eclipse. So far, so good. However, using the same laws and the same positions of the earth, moon, and sun, astronomers can retrodict the previous occurrence of a solar eclipse. The argument by which this retrodiction is made fulfills the requirements for a D-N explanation just as fully as does the prediction of the eclipse. Nevertheless, most of us would say that, while it is possible to explain an eclipse in terms of antecedent conditions, it is not possible to explain an eclipse in terms of subsequent conditions.

The moral: We invoke earlier conditions to explain subsequent facts; we do not invoke later conditions to explain earlier facts. The reason for this asymmetry seems to lie in the fact that causes, which have explanatory import, precede their effects—they do not follow their effects.

CE-4. The man and the pill. A man explains *his* failure to become pregnant during the past year on the ground that he has regularly consumed his wife's birth control pills, and that any man who regularly takes oral contraceptives will avoid getting pregnant.

The moral: This example shows that it is possible to construct valid deductive arguments with true premises in which some fact asserted by the premises is actually irrelevant. Since men do not get pregnant regardless, the fact that this man took birth control pills is irrelevant. Nevertheless, it conforms to the D-N pattern.

Counterexamples CE-1–CE-4 are all cases in which an argument that fulfills the Hempel-Oppenheim requirements manifestly fails to constitute a bona fide explanation. They were designed to show that these requirements are too weak to sort out the illegitimate explanations. A natural suggestion would be to strengthen them in ways that would rule out counterexamples of these kinds. For example, CE-1 and CE-2 could be disqualified if we stipulated that the antecedent conditions cited in the explanans must be causes of the explanandum. CE-3 could be eliminated by insisting that the so-called antecedent conditions must actually obtain prior to the explanandum. And CE-4 could be ruled out by stipulating that the antecedent conditions must be relevant to the explanandum. For various reasons Hempel declined to strengthen the requirements for D-N explanation in such ways.

The next counterexample has been offered as a case of a legitimate explanation that does not meet the Hempel-Oppenheim requirements.

CE-5. The ink stain. On the carpet, near the desk in Professor Jones's office, is an unsightly black stain. How does he explain it? Yesterday, an open bottle of black ink stood on his desk, near the corner. As he went by he accidentally bumped it with his elbow, and it fell to the floor, spilling ink on the carpet. This seems to be a perfectly adequate explanation; nevertheless, it does not incorporate any laws. Defenders of the D-N pattern would say that this is simply an incomplete explanation, and that the laws are tacitly assumed. Michael Scriven, who offered this example, argued that the explanation is clear and complete as it stands, and that any effort to spell out the laws and initial conditions precisely will meet with failure.

The moral: It is possible to have perfectly good explanations without any laws. The covering law conception is not universally correct.

The fifth counterexample raises profound problems concerning the nature of causality. Some philosophers, like Scriven, maintain that one event, such as the bumping of the ink bottle with the elbow, is obviously the cause of another event, such as the bottle falling off of the desk. Moreover, they claim, to identify the cause of an event is all that is needed to explain it. Other philosophers, including Hempel, maintain that a causal relation always involves (sometimes explicitly, sometimes implicitly) a general causal law. In the case of the ink stain, the relevant laws would include the laws of Newtonian mechanics (in explaining the bottle being knocked off the desk and falling to the floor) and some laws of chemistry (in explaining the stain on the carpet as a result of spilled ink).

1.11 TWO PATTERNS OF STATISTICAL EXPLANATION

Anyone who is familiar with any area of science—physical, biological, or social—realizes, as Hempel and Oppenheim had already noted, that not all explanations are of the deductive-nomological variety. Statistical laws play an important role in virtually every branch of contemporary science and statistical explanations—those falling into the two lower boxes in Table 1.1—are frequently given. In 1965b Hempel published a comprehensive essay, "Aspects of Scientific Explanation," in which he offered a theory of statistical explanation encompassing both types.

In the first type of statistical explanation, the *deductive-statistical* (*D-S*) *pattern*, statistical regularities are explained by deduction from more comprehensive statistical laws. Many examples can be found in contemporary science. For instance, archaeologists use the radiocarbon dating technique to ascertain the ages of pieces of wood or charcoal discovered in archaeological sites. If a piece of wood is found to have a concentration of C^{14} (a radioactive isotope of carbon) equal to one-fourth that of newly cut wood, it is inferred to be 11,460 years old. The reason is that the half-life of C^{14} is 5730 years, and in two half-lives it is extremely probable that about three-fourths of the C^{14} atoms will have decayed. Living trees replenish their supplies of C^{14} from the atmosphere; wood that has been cut cannot do so. Here is the D-S explanation:

(5) Every C^{14} atom (that is not exposed to external radiation[10]) has a probability of ½ of disintegrating within any period of 5730 years.

In any large collection of C^{14} atoms (that are not exposed to external radiation) approximately three-fourths will *very probably* decay within 11,460 years.

This derivation constitutes a deductive explanation of the probabilistic generalization that stands as its conclusion.

Deductive-statistical explanations are very similar, logically, to D-N explanations of generalizations. The only difference is that the explanation is a statistical law and the explanans must contain at least one statistical law. Universal laws have the form "All *A* are *B*" or "No *A* are *B*"; statistical laws say that a certain proportion of *A* are *B*.[11] Accordingly, the problem that plagued D-N explanations of universal generalizations also infects D-S explanations of statistical generalizations. Consider, for instance, one of the statistical generalizations in the preceding example—namely, that the half-life of C^{14} is 5730 years. There is a bona fide explanation of this generalization from the basic laws of quantum mechanics in conjunction with a description of the C^{14} nucleus. However, this statistical generalization can also be deduced from the conjunction of itself with Kepler's laws of planetary motion. This deduction would not qualify as any kind of legitimate explanation; like the case cited in Section 1.8, it would simply be a pointless derivation of the generalization about the half-life of C^{14} from itself.

Following the 1948 article, Hempel never returned to this problem concerning explanations of laws; he did not address it in Hempel (1965a), which contains characterizations of all four types of explanation represented in Table 1.1. This leaves both boxes on the right-hand side of Table 1.1 in a highly problematic status. Nevertheless, it seems clear that many sound explanations of both of these types can be found in the various sciences.

The second type of statistical explanation—the *inductive-statistical* (*I-S*) *pattern*—explains particular occurrences by subsuming them under statistical laws, much as D-N explanations subsume particular events under universal laws. Let us look at one of Hempel's famous examples. If we ask why Jane Jones recovered rapidly from her streptococcus infection, the answer is that she was given a dose of penicillin, and almost all strep infections clear up quickly upon administration of penicillin. More formally:

(6) Almost all cases of streptococcus infection clear up quickly after the administration of penicillin.
Jane Jones had a streptococcus infection.

[10] This qualification is required to assure that the disintegration is spontaneous and not induced by external radiation.

[11] As James Lennox remarks in Chapter 7 on philosophy of biology, Darwin's principle of natural selection is an example of a statistical law.

Jane Jones received treatment with penicillin.

$$\overline{\overline{\hspace{20em}}}\,[r]$$

Jane Jones recovered quickly.

This explanation is an argument that has three premises (the explanans); the first premise states a statistical regularity—a statistical law—while the other two state antecedent conditions. The conclusion (the explanandum) states the fact-to-be-explained. However, a crucial difference exists between explanations (3) and (6): D-N explanations subsume the events to be explained deductively, while I-S explanations subsume them inductively. The single line separating the premises from the conclusion in (3) signifies a relation of deductive entailment between the premises and conclusion. The double line in (6) represents a relationship of inductive support, and the attached variable *r* stands for the strength of that support. This strength of support may be expressed exactly, as a numerical value of a probability, or vaguely, by means of such phrases as "very probably" or "almost certainly."

An explanation of either of these two kinds 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; in an I-S explanation the event to be explained has high inductive probability relative to the explanatory facts. This feature of expectability is closely related to the *explanation-prediction symmetry thesis* for explanations of particular facts. According to this thesis any acceptable explanation of a particular fact is an argument, deductive or inductive, that could have been used to predict the fact in question if the facts stated in the explanans had been available prior to its occurrence.[12] As we shall see, this symmetry thesis met with serious opposition.

Hempel was not by any means the only philosopher in the early 1960s to notice that statistical explanations play a significant role in modern science. He was, however, the first to present a detailed account of the nature of statistical explanation, and the first to bring out a fundamental problem concerning statistical explanations of particular facts. The case of Jane Jones and her quick recovery can be used as an illustration. It is well known that certain strains of the streptococcus bacterium are penicillin-resistant, and if Jones's infection were of that type, the probability of her quick recovery after treatment with penicillin would be small. We could, in fact, set up the following inductive argument:

(7) Almost no cases of penicillin-resistant streptococcus infection clear up quickly after the administration of penicillin.
Jane Jones had a penicillin-resistant streptococcus infection.
Jane Jones received treatment with penicillin.

$$\overline{\overline{\hspace{20em}}}\,[q]$$

Jane Jones did not recover quickly.

The remarkable fact about arguments (6) and (7) is that their premises are mutually compatible—they could all be true. Nevertheless, their conclusions contra-

[12] This thesis was advanced for D-N explanation in Hempel-Oppenheim (1948, 249), and reiterated, with some qualifications, for D-N and I-S explanations in Hempel (1965a, Sections 2.4, 3.5).

dict one another. This is a situation that can never occur with deductive arguments. Given two valid deductions with incompatible conclusions, their premises must also be incompatible. Thus, the problem that has arisen in connection with I-S explanations has no analog in D-N explanations. Hempel called this *the problem of ambiguity of I-S explanation.*

The source of the problem of ambiguity is a simple and fundamental difference between universal laws and statistical laws. Given the proposition that all *A* are *B*, it follows immediately that all things that are both *A* and *C* are *B*. If all humans are mortal, then all people who are over six feet tall are mortal. However, even if almost all humans who are alive now will be alive five years from now, *it does not follow* that almost all living humans with advanced cases of pancreatic cancer will be alive five years hence. As we noted in Section 1.5, there is a parallel fact about arguments. Given a valid deductive argument, the argument will remain valid if additional premises are supplied, as long as none of the original premises is taken away. Deduction is erosion-proof. Given a strong inductive argument—one that supports its conclusion with a high degree of probability—the addition of one more premise may undermine it completely. For centuries Europeans had a great body of inductive evidence to support the proposition that all swans are white, but one true report of a black swan in Australia completely refuted that conclusion. Induction is not erosion-proof.

Hempel sought to resolve the problem of ambiguity by means of his *requirement of maximal specificity* (*RMS*). It is extremely tricky to state RMS with precision, but the basic idea is fairly simple. In constructing I-S explanations we must include all relevant knowledge we have that would have been available, in principle, prior to the explanandum-fact. If the information that Jones's infection is of the penicillin-resistant type is available to us, argument (6) would not qualify as an acceptable I-S explanation.[13]

In Section 1.8 we stated Hempel and Oppenheim's four conditions of adequacy for D-N explanations. We can now generalize these conditions so that they apply both to D-N and I-S explanations as follows:

1. The explanation must be an argument having correct (deductive or inductive) logical form.
2. The explanans must contain at least one general law (universal or statistical), and this law must actually be required for the derivation of the explanandum.
3. The explanans must have empirical content; it must be capable, at least in principle, of test by experiment or observation.
4. The sentences constituting the explanans must be true.
5. The explanation must satisfy the requirement of maximal specificity.[14]

[13] Nor would (6) qualify as an acceptable I-S explanation if we had found that Jones's infection was of the non-penicillin-resistant variety, for the probability of quick recovery among people with that type of infection is different from the probability of quick recovery among those who have an unspecified type of streptococcus infection.

[14] D-N explanations of particular facts automatically satisfy this requirement. If all *A* are *B*, the probability that an *A* is a *B* is one. Under those circumstances, the probability that an *A* which is also a *C* is a *B* is also one. Therefore, no partition of *A* is relevant to *B*.

The theory of scientific explanation developed by Hempel in his "Aspects" essay won rather wide approval among philosophers of science. During the mid-to-late 1960s and early 1970s it could appropriately be considered *the received view of scientific explanation*. According to this view, every legitimate scientific explanation must fit the pattern corresponding to one or another of the four boxes in Table 1.1.

1.12 CRITICISMS OF THE I-S PATTERN OF SCIENTIFIC EXPLANATION

We noticed in Section 1.10 that major criticisms of the D-N pattern of scientific explanation can be posed by means of well-known counterexamples. The same situation arises in connection with the I-S pattern. Consider the following:

CE-6. Psychotherapy. Suppose that Bruce Brown has a troublesome neurotic symptom. He undergoes psychotherapy and his symptom disappears. Can we explain his recovery in terms of the treatment he has undergone? We could set out the following inductive argument, in analogy with argument (6):

(8) Most people who have a neurotic symptom of type *N* and who undergo psychotherapy experience relief from that symptom.
Bruce Brown had a symptom of type *N* and he underwent psychotherapy.
══════════════════════════════ [*r*]
Bruce Brown experienced relief from his symptom.

Before attempting to evaluate this proffered explanation we should take account of the fact that there is a fairly high spontaneous remission rate—that is, many people who suffer from that sort of symptom get better regardless of treatment. No matter how large the number *r*, if the rate of recovery for people who undergo psychotherapy is no larger than the spontaneous remission rate, it would be a mistake to consider argument (8) a legitimate explanation. A high probability is not *sufficient* for a correct explanation. If, however, the number *r* is not very large, but is greater than the spontaneous remission rate, the fact that the patient underwent psychotherapy has at least some degree of explanatory force. A high probability is not *necessary* for a sound explanation.

Another example reinforces the same point.

CE-7. Vitamin C and the common cold.[15] Suppose someone were to claim that large doses of vitamin C would produce rapid cures for the common cold. To ascertain the efficacy of vitamin C in producing rapid recovery from colds, we should note, it is *not* sufficient to establish that most people recover quickly; most colds disappear within a few days regardless of treatment. What is required is a double-

[15] Around the time Hempel was working out his theory of I-S explanation, Linus Pauling's claims about the value of massive doses of vitamin C in the prevention of common colds was receiving a great deal of attention. Although Pauling made no claims about the ability of vitamin C to cure colds, it occurred to me that a fictitious example of this sort could be concocted.

blind controlled experiment[16] in which the rate of quick recovery for those who take vitamin C is compared with the rate of quick recovery for those who receive only a placebo. If there is a significant difference in the probability of quick recovery for those who take vitamin C and for those who do not, we may conclude that vitamin C has some degree of causal efficacy in lessening the duration of colds. If, however, there is no difference between the two groups, then it would be a mistake to try to explain a person's quick recovery from a cold by constructing an argument analogous to (6) in which that result is attributed to treatment with vitamin C.

The moral: CE-6 and CE-7 call attention to the same point as CE-4 (the man and the pill). All of them show that something must be done to exclude irrelevancies from scientific explanations. If the rate of pregnancy among men who consume oral contraceptives is the same as for men who do not, then the use of birth control pills is causally and explanatorily irrelevant to pregnancy among males. Likewise, if the rate of relief from neurotic symptoms is the same for those who undergo psychotherapy as it is for those who do not, then psychotherapy is causally and explanatorily irrelevant to the relief from neurotic symptoms. Again, if the rate of rapid recovery from common colds is the same for those who do and those who do not take massive doses of vitamin C, then consumption of massive doses of vitamin C is causally and explanatorily irrelevant to rapid recovery from colds.[17] Hempel's requirement of maximal specificity was designed to insure that *all* relevant information (of a suitable sort) is included in I-S explanations. What is needed in addition is a requirement insuring that *only* relevant information is included in D-N *or* I-S explanations.

CE-8. Syphilis and paresis. Paresis is a form of tertiary syphilis which can be contracted only by people who go through the primary, secondary, and latent forms of syphilis without treatment with penicillin. If one should ask why a particular person suffers from paresis, a correct answer is that he or she was a victim of untreated latent syphilis. Nevertheless, only a small proportion of those with untreated latent syphilis—about 25%—actually contract paresis. Given a randomly selected member of the class of victims of untreated latent syphilis, one should predict that that person *will not* develop paresis.

The moral: there are legitimate I-S explanations in which the explanans *does not* render the explanandum highly probable. CE-8 responds to the explanation-prediction symmetry thesis—the claim that an explanation is an argument of such a sort that it could have been used to predict the explanandum if it had been available prior to the

[16] In a controlled experiment there are two groups of subjects, the experimental group and the control group. These groups should be as similar to one another as possible. The members of the experimental group receive the substance being tested, vitamin C. The members of the control group receive a placebo, that is, an inert substance such as a sugar pill that is known to have no effect on the common cold. In a blind experiment the subjects do not know whether they are receiving vitamin C or the placebo. This is important, for if the subjects knew which treatment they were receiving, the power of suggestion might skew the results. An experiment is double-blind if neither the person who hands out the pills nor the subjects know which subject is getting which type of pill. If the experiment is not double-blind, the person administering the pills might, in spite of every effort not to, convey some hint to the subject.

[17] It should be carefully noted that I am claiming *neither* that psychotherapy is irrelevant to remission of neurotic symptoms *nor* that vitamin C is irrelevant to rate of recovery from colds. I *am* saying that that is the point at issue so far as I-S explanation is concerned.

fact-to-be-explained. It is worth noting, in relation to CE-6 and CE-7, that untreated latent syphilis is highly relevant to the occurrence of paresis, although it does not make paresis highly probable, or even more probable than not.

CE-9. The biased coin. Suppose that a coin is being tossed, and that it is highly biased for heads—in fact, on any given toss, the probability of getting heads is 0.95, while the probability of tails is 0.05. The coin is tossed and comes up heads. We can readily construct an I-S explanation fitting all of the requirements. But suppose it comes up tails. In this case an I-S explanation is out of the question. Nevertheless, to the degree that we understand the mechanism involved, and consequently the probable outcome of heads, to that same degree we understand the improbable outcome, even though it occurs less frequently.

The moral: If we are in a position to construct statistical explanations of events that are highly probable, then we also possess the capability of framing statistical explanations of events that are extremely improbable.

1.13 DETERMINISM, INDETERMINISM, AND STATISTICAL EXPLANATION

When we look at an I-S explanation such as (6), there is a strong temptation to regard it as incomplete. It may, to be sure, incorporate all of the relevant knowledge we happen to possess. Nevertheless, we may feel, it is altogether possible that medical science will discover enough about streptococcus infections and about penicillin treatment to be able to determine precisely which individuals with strep infections will recover quickly upon treatment with penicillin and which individuals will not. When that degree of knowledge is available we will not have to settle for I-S explanations of rapid recoveries from strep infections; we will be able to provide D-N explanations instead. Similar remarks can also be made about several of the counterexamples—in particular, examples CE-6–CE-9.

Consider CE-8, the syphilis-paresis example. As remarked above, with our present state of knowledge we can predict that about 25% of all victims of untreated latent syphilis contract paresis, but we do not know how to distinguish those who will develop paresis from those who will not. Suppose Sam Smith develops paresis. At this stage of our knowledge the best we can do by way of an I-S explanation of Smith's paresis is the following:

(9) 25% of all victims of untreated latent syphilis develop paresis.
Smith had untreated latent syphilis.
══════════════════════ [.25]
Smith contracted paresis.

This could not be accepted as an I-S explanation because of the weakness of the relation of inductive support.

Suppose that further research on the causes of paresis reveals a factor in the blood—call it the *P*-factor—which enables us to pick out, with fair reliability—say

95%—those who will develop paresis. Given that Smith has the *P*-factor, we can construct the following argument:

(10) 95% of all victims of untreated latent syphilis who have the *P*-factor develop paresis.
Smith had untreated latent syphilis.
Smith had the *P*-factor.
═══════════════════════════ [.95]
Smith developed paresis.

In the knowledge situation just described, this would count as a pretty good I-S explanation, for 0.95 is fairly close to 1.

Let us now suppose further that additional medical research reveals that, among those victims of untreated latent syphilis who have the *P*-factor, those whose spinal fluid contains another factor *Q* invariably develop paresis. Given that information, and the fact that Smith has the *Q*-factor, we can set up the following explanation:

(11) All victims of untreated latent syphilis who have the *P*-factor and the *Q*-factor develop paresis.
Smith had untreated latent syphilis.
Smith had the *P*-factor.
Smith had the *Q*-factor.
───────────────────────────
Smith developed paresis.

If the suppositions about the *P*-factor and the *Q*-factor were true, this argument would qualify as a correct D-N explanation. We accepted (10) as a correct explanation of Smith's paresis only because we were lacking the information that enabled us to set up (11).

Determinism is the doctrine that says that everything that happens in our universe is completely determined by prior conditions.[18] If this thesis is correct, then each and every event in the history of the universe—past, present, or future—is, in principle, deductively explainable. If determinism is true, then every sound I-S explanation is merely an incomplete D-N explanation. Under these circumstances, the I-S pattern is not really a stand-alone type of explanation; all fully correct explanations fit the D-N pattern. The lower left-hand box of Table 1.1 would be empty. This does *not* mean that I-S explanations—that is, incomplete D-N explanations—are useless, only that they are incomplete.

Is determinism true? We will not take a stand on that issue in this chapter. Modern physics—quantum mechanics in particular—seems to offer strong reasons to believe that determinism is false, but not everyone agrees with this interpretation. However, we will take the position that indeterminism *may* be true, and see what the consequences are with respect to statistical explanation.

According to most physicists and philosophers of physics, the spontaneous disintegration of the nucleus of an atom of a radioactive substance is a genuinely indeterministic happening. Radioactive decay is governed by laws, but they are

[18] Determinism is discussed in detail in Chapter 6.

fundamentally and irreducibly statistical. Any C^{14} atom has a fifty-fifty chance of spontaneously disintegrating within the next 5730 years and a fifty-fifty chance of not doing so. Given a collection of C^{14} atoms, the probability is overwhelming that some will decay and some will not in the next 5730 years. However, no way exists, even in principle, to select in advance those that will. No D-N explanation of the decay of any such atom can possibly be constructed; however, I-S explanations can be formulated. For example, in a sample containing 1 milligram of C^{14} there are approximately 4×10^{19} atoms. If, in a period of 5730 years, precisely half of them decayed, approximately 2×10^{19} would remain intact. It is *extremely unlikely* that *exactly* half of them would disintegrate in that period, but it is *extremely likely* that *approximately* half would decay. The following argument—which differs from (5) by referring to one particular sample *S*—would be a strong I-S explanation:

(12) *S* is a sample of C^{14} that contained one milligram 5730 years ago.
S has not been exposed to external radiation.[19]
The half-life of C^{14} is 5730 years.
$\overline{\overline{\hspace{20em}}}[r]$
S now contains one-half milligram ($\pm 1\%$) of C^{14}.

In this example, *r* differs from 1 by an incredibly tiny margin, but is not literally equal to 1. In a world that is not deterministic, I-S explanations that are not merely incomplete D-N explanations can be formulated.

1.14 THE STATISTICAL RELEVANCE (S-R) MODEL OF EXPLANATION

According to the received view, scientific explanations are arguments; each type of explanation in Table 1.1 is some type of argument satisfying certain conditions. For this reason, we can classify the received view as an *inferential conception* of scientific explanation. Because of certain difficulties, associated primarily with I-S explanation, another pattern for statistical explanations of particular occurrences was developed. A fundamental feature of this model of explanation is that it *does not* construe explanations as arguments.

One of the earliest objections to the I-S pattern of explanation—as shown by CE-6 (psychotherapy) and CE-7 (vitamin C and the common cold)—is that statistical relevance rather than high probability is the crucial relationship in statistical explanations. Statistical relevance involves a relationship between two different probabilities. Consider the psychotherapy example. Bruce Brown is a member of the class of people who have a neurotic symptom of type *N*. Within that class, regardless of what the person does in the way of treatment or nontreatment, there is a certain probability of relief from the symptom (*R*). That is the *prior probability* of recovery; let us symbolize it as "*Pr(R/N)*." Then there is a probability of recovery in the class of

[19] This qualification is required to assure that the disintegrations have been spontaneous and not induced by external radiation.

people with that symptom who undergo psychotherapy (P); it can be symbolized as "$Pr(R/N . P)$." If

$$Pr(R/N . P) > Pr(R/N)$$

then psychotherapy is positively relevant to recovery, and if

$$Pr(R/N . P) < Pr(R/N)$$

then psychotherapy is negatively relevant to recovery. If

$$Pr(R/N . P) = Pr(R/N)$$

then psychotherapy is irrelevant to recovery. Suppose psychotherapy is positively relevant to recovery. If someone then asks why Bruce Brown, who suffered with neurotic symptom N, recovered from his symptom, we can say that it was because he underwent psychotherapy. That is at least an important part of the explanation.

Consider another example. Suppose that Grace Green, an American woman, suffered a serious heart attack. In order to explain why this happened we search for factors that are relevant to serious heart attacks—for example, smoking, high cholesterol level, and body weight. If we find that she was a heavy cigarette smoker, had a serum cholesterol level above 300, and was seriously overweight, we have at least a good part of an explanation, for all of those factors are positively relevant to serious heart attacks. There are, of course, other relevant factors, but these three will do for purposes of illustration.

More formally, if we ask why this member of the class A (American women) has characteristic H (serious heart attack), we can take the original reference class A and subdivide or *partition* it in terms of such relevant factors as we have mentioned: S (heavy cigarette smokers), C (high cholesterol level), and W (overweight). This will give us a partition with eight cells (where the dot signifies conjunction and the tilde "~" signifies negation):

$S . C . W$	$\sim S . C . W$
$S . C . \sim W$	$\sim S . C . \sim W$
$S . \sim C . W$	$\sim S . \sim C . W$
$S . \sim C . \sim W$	$\sim S . \sim C . \sim W$

An S-R explanation of Green's heart attack has three parts:

1. The prior probability of H, namely, $Pr(H/A)$.
2. The posterior probabilities of H with respect to each of the eight cells, $Pr(H/S . C . W)$, $Pr(H/S . C . \sim W)$, . . . , $Pr(H/ \sim S . \sim C . \sim W)$.
3. The statement that Green is a member of $S . C . W$.

It is stipulated that the partition of the reference class must be made in terms of all and only the factors relevant to serious heart attacks.

Clearly, an explanation of that sort is not an argument; it has neither premises nor conclusion. It does, of course, consist of an explanans and an explanandum.

Items 1–3 constitute the explanans; the explanandum is Green's heart attack. Moreover, no restrictions are placed on the size of the probabilities—they can be high, middling, or low. All that is required is that these probabilities differ from one another in various ways, because we are centrally concerned with relations of statistical relevance.

Although the S-R pattern of scientific explanation provides some improvements over the I-S model, it suffers from a fundamental inadequacy. It focuses on statistical relevance rather than causal relevance. It may, as a result, tend to foster a confusion of causes and correlations. In the vitamin C example, for instance, we want a controlled experiment to find out whether taking massive doses of vitamin C is *causally relevant* to quick recovery from colds. We attempt to find out whether taking vitamin C is *statistically relevant* to rapid relief because the statistical relevance relation is evidence regarding the presence or absence of *causal relevance*. It is causal relevance that has genuine explanatory import. The same remark applies to other examples as well. In the psychotherapy example we try to find out whether such treatment is statistically relevant to relief from neurotic symptoms in order to tell whether it is causally relevant. In the case of the heart attack, many clinical studies have tried to find statistical relevance relations as a basis for determining what is causally relevant to the occurrence of serious heart attacks.

1.15 TWO GRAND TRADITIONS

We have been looking at the development of the received view, and at some of the criticisms that have been leveled against it. The strongest intuitive appeal of that view comes much more from explanations of laws than from explanations of particular facts. One great example is the *Newtonian synthesis*. Prior to Newton we had a miscellaneous collection of laws including Kepler's three laws of planetary motion and Galileo's laws of falling objects, inertia, projectile motion, and pendulums. By invoking three simple laws of motion and one law of gravitation, Newton was able to explain these laws—and in some cases correct them. In addition, he was able to explain many other regularities, such as the behavior of comets and tides, as well. Later on, the molecular-kinetic theory provided a Newtonian explanation of many laws pertaining to gases. Quite possibly the most important feature of the Newtonian synthesis was the extent to which it systematized our knowledge of the physical world by subsuming all sorts of regularities under a small number of very simple laws. Another excellent historical example is the explanation of light by subsumption under Maxwell's theory of electromagnetic radiation.

The watchword in these beautiful historical examples is *unification*. A large number of specific regularities are unified in one theory with a small number of assumptions or postulates. This theme was elaborated by Michael Friedman (1974) who asserted that our comprehension of the universe is increased as the number of independently acceptable assumptions we require is reduced. I would be inclined to add that this sort of systematic unification of our scientific knowledge provides a comprehensive world picture or worldview. This, I think, represents one major aspect of scientific explanation—it is the notion that we understand what goes on in the world if we can fit it into a comprehensive worldview. As Friedman points out, this is a *global* conception of explanation. The value of explanation lies in fitting things into a universal pattern, or a pattern that covers major segments of the universe.[20]

As we look at many of the criticisms that have been directed against the received view, it becomes clear that causality is a major focus. Scriven offered his ink stain example, CE-5, to support the claim that finding the explanation amounts, in many cases, simply to finding the causes. This is clearly explanation on a very *local* level. All we need to do, according to Scriven, is to get a handle on events in an extremely limited spacetime region that led up, causally, to the stain on the carpet, and we have adequate understanding of that particular fact. In this connection, we should also recall CE-1 and CE-2. In the first of these we sought a local causal explanation for the length of a shadow, and in the second we wanted a causal explanation for a particular storm. Closely related noncausal ''explanations'' were patently unacceptable. In such cases as the Chernobyl accident and the Challenger space-shuttle explosion we also seek causal explanations, partly in order to try to avoid such tragedies in the future. Scientific explanation has its practical as well as its purely intellectual value.

It often happens, when we try to find causal explanations for various occurrences, that we have to appeal to entities that are not directly observable with the unaided human senses. For example, to understand AIDS (Acquired Immunodeficiency Syndrome), we must deal with viruses and cells. To understand the transmission of traits from parents to offspring, we become involved with the structure of the DNA molecule. To explain a large range of phenomena associated with the nuclear accident at Three Mile Island, we must deal with atoms and subatomic particles. When we try to construct causal explanations we are attempting to discover the mechanisms—often hidden mechanisms—that bring about the facts we seek to understand. The search for causal explanations, and the associated attempt to expose the hidden workings of nature, represent a second grand tradition regarding scientific explanation. We can refer to it as *the causal-mechanical tradition*.

Having contrasted the two major traditions, we should call attention to an important respect in which they overlap. When the search for hidden mechanisms is successful, the result is often to reveal a small number of basic mechanisms that underlie wide ranges of phenomena. The explanation of diverse phenomena in terms of the same mechanisms constitutes theoretical unification. For instance, the kinetic-molecular theory of gases unified thermodynamic phenomena with Newtonian particle mechanics. The discovery of the double-helical structure of DNA, for another example, produced a major unification of biology and chemistry.

Each of the two grand traditions faces certain fundamental problems. The tradition of explanation as unification—associated with the received view—still faces the problem concerning explanations of laws that was pointed out in 1948 by Hempel and Oppenheim. It was never solved by Hempel in any of his subsequent work on scientific explanation. If the technical details of Friedman's theory of unification were satisfactory, it would provide a solution to that problem. Unfortunately, it appears to encounter serious technical difficulties (see Kitcher 1976 and Salmon 1989).

[20] The unification approach has been dramatically extended and improved by Philip Kitcher (1976, 1981, and 1989).

The causal-mechanical tradition faces a longstanding philosophical difficulty concerning the nature of causality that had been posed by David Hume in the eighteenth century. The problem—stated extremely concisely—is that we seem unable to identify the *connection* between cause and effect, or to find the *secret power* by which the cause brings about the effect. Hume is able to find certain *constant conjunctions*—for instance, between fire and heat—but he is unable to find the connection. He is able to see the spatial contiguity of events we identify as cause and effect, and the temporal priority of the cause to the effect—as in collisions of billiard balls, for instance—but still no *necessary connection*. In the end he locates the connection in the human imagination—in the psychological expectation we feel with regard to the effect when we observe the cause.[21]

Hume's problem regarding causality is one of the most recalcitrant in the whole history of philosophy. Some philosophers of science have tried to provide a more objective and robust concept of causality, but none has enjoyed widespread success. One of the main reasons the received view was reticent about incorporating causal considerations in the analysis of scientific explanation was an acute sense of uneasiness about Hume's problem. One of the weaknesses of the causal view, as it is handled by many philosophers who espouse it, is the absence of any satisfactory theory of causality.[22]

1.16 THE PRAGMATICS OF EXPLANATION

As we noted in Section 1.4, the term "explanation" refers sometimes to linguistic entities—that is, collections of statements of facts—and sometimes to nonlinguistic entities—namely, those very facts. When we think in terms of the human activity of explaining something to some person or group of people, we are considering linguistic behavior. Explaining something to someone involves uttering or writing statements. In this section we look at some aspects of this *process of explaining*. In this chapter, up to this point, we have dealt mainly with the *product* resulting from this activity, that is, the explanation that was offered in the process of explaining.

When philosophers discuss language they customarily divide the study into three parts: syntax, semantics, and pragmatics. Syntax is concerned only with relationships among the symbols, without reference to the meanings of the symbols or the people who use them. Roughly speaking, syntax is pure grammar; it deals with the conventions governing combinations and manipulations of symbols. Semantics deals with the relationships between symbols and the things to which the symbols refer. Meaning and truth are the major semantical concepts. Pragmatics deals with the relationships among symbols, what they refer to, and the users of language. Of particular interest for our discussion is the treatment of the context in which language is used.

The 1948 Hempel-Oppenheim essay offered a highly formalized account of D-N explanations of particular facts, and it characterized such explanations in syntactical

[21] Hume's analysis of causation is discussed in greater detail in Chapter 2, Part II.

[22] I have tried to make some progress in this direction in Salmon (1984, Chapters 5–7).

and semantical terms alone. Pragmatic considerations were not dealt with. Hempel's later characterization of the other types of explanations were given mainly in syntactical and semantical terms, although I-S explanations are, as we noted, relativized to knowledge situations. Knowledge situations are aspects of the human contexts in which explanations are sought and given. Such contexts have other aspects as well.

One way to look at the pragmatic dimensions of explanation is to start with the question by which an explanation is sought. In Section 1.3 we touched briefly on this matter. We noted that many, if not all, explanations can properly be requested by *explanation-seeking why-questions*. In many cases, the first pragmatic step is to clarify the question being asked; often the sentence uttered by the questioner depends upon contextual clues for its interpretation. As Bas van Fraassen, one of the most important contributors to the study of the pragmatics of explanation, has shown, the emphasis with which a speaker poses a question may play a crucial role in determining just what question is being asked. He goes to the Biblical story of the Garden of Eden to illustrate. Consider the following three questions:

(i) Why did Adam eat *the apple?*
(ii) Why did *Adam* eat the apple?
(iii) Why did Adam *eat* the apple?

Although the words are the same—and in the same order—in each, they pose three very different questions. This can be shown by considering what van Fraassen calls the *contrast class*. Sentence (i) asks why Adam ate the apple instead of a pear, a banana, or a pomegranate. Sentence (ii) asks why Adam, instead of Eve, the serpent, or a goat, ate the apple. Sentence (iii) asks why Adam ate the apple instead of throwing it away, feeding it to a goat, or hiding it somewhere. Unless we become clear on the question being asked, we can hardly expect to furnish appropriate answers.

Another pragmatic feature of explanation concerns the knowledge and intellectual ability of the person or group requesting the explanation. On the one hand, there is usually no point in including in an explanation matters that are obvious to all concerned. Returning to (3)—our prime example of a D-N explanation of a particular fact—one person requesting an explanation of the sudden dramatic increase in the skater's rate of rotation might have been well aware of the fact that she drew her arms in close to her body, but unfamiliar with the law of conservation of angular momentum. For this questioner, knowledge of the law of conservation of angular momentum is required in order to understand the explanandum-fact. Another person might have been fully aware of the law of conservation of angular momentum, but failed to notice what the skater did with her arms. This person needs to be informed of the skater's arm maneuver. Still another person might have noticed the arm maneuver, and might also be aware of the law of conservation of angular momentum, but failed to notice that this law applies to the skater's movement. This person needs to be shown how to apply the law in the case in question.

On the other hand, there is no point in including material in an explanation that is beyond the listeners' ability to comprehend. To most schoolchildren, for example, an explanation of the darkness of the night sky that made reference to the non-Euclidean structure of space or the mean free path of a photon would be inappropriate. Many of the explanations we encounter in real-life situations are incomplete on account of the explainer's view of the background knowledge of the audience.

A further pragmatic consideration concerns the interests of the audience. A scientist giving an explanation of a serious accident to a congressional investigating committee may tell the members of Congress far more than they want to know about the scientific details. In learning why an airplane crashed, the committee might be very interested to find that it was because of an accumulation of ice on the wing, but totally bored by the scientific reason why ice-accumulations cause airplanes to crash.

Peter Railton (1981) has offered a distinction that helps considerably in understanding the role of pragmatics in scientific explanation. First, he introduces the notion of an *ideal explanatory text.* An ideal explanatory text contains *all* of the facts and *all* of the laws that are relevant to the explanandum-fact. It details *all* of the causal connections among those facts and *all* of the hidden mechanisms. In most cases the ideal explanatory text is huge and complex. Consider, for example, an explanation of an automobile accident. The *full* details of such items as the behavior of both drivers, the operations of both autos, the condition of the highway surface, the dirt on the windshields, and the weather, would be unbelievably complicated. That does not really matter, for the ideal explanatory text is seldom, if ever, spelled out fully. What is important is to have the ability to illuminate portions of the ideal text as they are wanted or needed. When we do provide knowledge to fill in some aspect of the ideal text we are furnishing *explanatory information.*

A request for a scientific explanation of a given fact is almost always—if not literally always—a request, not for the ideal explanatory text, but for explanatory information. The ideal text contains all of the facts and laws pertaining to the explanandum-fact. These are the completely objective and nonpragmatic aspects of the explanation. If explanatory information is to count as legitimate it must correspond to the objective features of the ideal text. The ideal text determines what is *relevant* to the explanandum-fact. Since, however, we cannot provide the whole ideal text, nor do we want to, a selection of information to be supplied must be made. This depends on the knowledge and interests of those requesting and those furnishing explanations. The information that satisfies the request in terms of the interests and knowledge of the audience is *salient* information. The pragmatics of explanation determines salience—that is, what aspects of the ideal explanatory text are appropriate for an explanation in a particular context.

1.17 CONCLUSION

Several years ago, a friend and colleague—whom I will call *the friendly physicist*—was sitting on a jet airplane awaiting takeoff. Directly across the aisle was a young boy holding a helium-filled balloon by a string. In an effort to pique the child's curiosity, the friendly physicist asked him what he thought the balloon would do when the plane accelerated for takeoff. After a moment's thought the boy said that it would move toward the back of the plane. The friendly physicist replied that *he* thought it would move toward the front of the cabin. Several adults in the vicinity became interested in the conversation, and they insisted that the friendly physicist was wrong. A flight attendant offered to wager a miniature bottle of Scotch that he was mistaken—a bet that he was quite willing to accept. Soon thereafter the plane accelerated, the balloon moved forward, and the friendly physicist enjoyed a free drink.[23]

Why did the balloon move toward the front of the cabin? Two explanations can be offered, both of which are correct. First, one can tell a story about the behavior of the molecules that made up the air in the cabin, explaining how the rear wall collided with nearby molecules when it began its forward motion, thus creating a pressure gradient from the back to the front of the cabin. This pressure gradient imposed an unbalanced force on the back side of the balloon, causing it to move forward with respect to the walls of the cabin.[24] Second, one can cite an extremely general physical principle—Einstein's *principle of equivalence*—according to which an acceleration is physically equivalent, from the standpoint of the occupants of the cabin, to a gravitational field. Since helium-filled balloons tend to rise in the atmosphere in the earth's gravitational field, they will move forward when the airplane accelerates, reacting just as they would if a massive object were suddenly placed behind the rear wall.

The first of these explanations is causal-mechanical. It appeals to unobservable entities, describing the causal processes and causal interactions involved in the explanandum phenomenon. When we are made aware of these explanatory facts we understand how the phenomenon came about. This is the kind of explanation that advocates of the causal-mechanical tradition find congenial. The second explanation illustrates the unification approach. By appealing to an extremely general physical principle, it shows how this odd little occurrence fits into the universal scheme of things. It does not refer to the detailed mechanisms. This explanation provides a different kind of understanding of the same fact.

Which of these explanations is correct? Both are. Both of them are embedded in the ideal explanatory text. Each of them furnishes valuable explanatory information. It would be a serious error to suppose that any phenomenon has only one explanation. It is a mistake, I believe, to ask for *the* explanation of any occurrence. Each of these explanations confers a kind of scientific understanding. Pragmatic considerations might dictate the choice of one rather than the other in a given context. For example, the explanation in terms of the equivalence principle would be unsuitable for a ten-year-old child. The same explanation might be just right in an undergraduate physics course. But both are bona fide explanations.

As we noted in Section 1.10, the 1948 Hempel-Oppenheim essay attracted almost no attention for about a decade after its publication. Around 1959 it became the focus of intense controversy, much of it stemming from those who saw causality as central to scientific explanation. The subsequent thirty years have seen a strong opposition between the advocates of the received view and the proponents of causal explanation. Each of the two major approaches has evolved considerably during this

[23] This little story was previously published in Salmon (1980). I did not offer an explanation of the phenomenon in that article.

[24] Objects that are denser than air do not move toward the front of the cabin because the pressure difference is insufficient to overcome their inertia.

period—indeed, they have developed to the point that they can peacefully coexist as two distinct aspects of scientific explanation. Scientific understanding is, after all, a complicated affair; we should not be surprised to learn that it has many different aspects. Exposing underlying mechanisms and fitting phenomena into comprehensive pictures of the world seem to constitute two important aspects. Moreover, as remarked above, we should remember that these two types of understanding frequently overlap. When we find that the same mechanisms underlie diverse types of natural phenomena this ipso facto constitutes a theoretical unification.

On one basic thesis there is nearly complete consensus. Recall that in the early decades of the twentieth century many scientists and philosophers denied that there can be any such thing as scientific explanation. Explanation is to be found, according to this view, only in the realms of theology and metaphysics. At present it seems virtually unanimously agreed that, however it may be explicated, there is such a thing as scientific explanation. Science *can* provide deep understanding of our world. We do *not* need to appeal to supernatural agencies to achieve understanding. Equally importantly, we can contrast the objectively based explanations of contemporary science with the pseudounderstanding offered by such flagrantly unscientific approaches as astrology, creation science, and scientology. These are points worth remembering in an age of rampant pseudoscience.

QUESTIONS

1. Must every scientific explanation contain a law of nature? According to philosophers who support "the received view" the answer is affirmative. Other philosophers have answered in the negative. Discuss critically the arguments pro and con. Give your own answer, supported by reasons.
2. Are there any inductive or statistical explanations of particular facts? In their classic 1948 paper Hempel and Oppenheim say that there are such explanations, but do not offer any explication of their nature. Later attempts to work out the details ran into many difficulties. Discuss these problems and say whether you think they are insuperable. Give your reasons.
3. According to the *explanation-prediction symmetry thesis*, every satisfactory scientific explanation could (in some suitable context) serve as a scientific prediction, and every scientific prediction could (in some suitable context) serve as a scientific explanation. Critically discuss both parts of this symmetry thesis. Give your reasons for accepting or rejecting each part.
4. Are there any *fundamental* differences between explanations in the natural sciences and explanations in the social sciences? (See Merrilee H. Salmon's chapter on philosophy of the social sciences.) Are there basic differences between human behavior and the behavior of other kinds of physical objects that make one kind more amenable to explanation than the other? Is explanation of human behavior that involves conscious deliberation and free choice possible? Discuss critically.
5. In this chapter it was suggested that "No gold sphere has a mass greater than 100,000 kg" is not a lawlike statement, whereas "No enriched uranium sphere has a mass greater than 100,000 kg" is a lawlike statement. Discuss the distinction between lawlike and accidental generalizations. Explain as clearly as possible why one is lawlike and the other is not.
6. Discuss the role of causality in scientific explanation. Do all legitimate scientific explanations make reference to causal relations? Is causality essentially irrelevant to scientific explanation? Are some good explanations causal and other good explanations not? Discuss critically.

7. Choose an actual example of a scientific explanation from a magazine such as *Scientific American, Science, Nature, American Scientist,* or from a textbook you have used in a science course. Give a concise summary of this explanation, and analyze it in terms of the models (such as D-N, I-S, D-S, S-R) and concepts (such as covering law, causal-mechanical, unification) introduced in this chapter. Evaluate the explanation in terms of these models and/or concepts.
8. In Section 1.9 it was claimed that

   (i) All gases, kept in closed containers of fixed size, exert greater pressure when heated

   is a general statement, whereas

   (v) All Apache basketry is made by women

   is not completely general because it refers specifically to a particular group of people. But, it might be objected, (i) refers to physical objects of a specific type, namely, gases in closed containers, so it is not completely general either. Moreover, (v) is a general statement about the Apache. Discuss this objection. Hint: Statement (i) can be formulated as follows: "If *anything* is a gas in a closed container that is heated, it will expand." But: Statement (v) can likewise be reformulated as follows: "If anything is an Apache basket, it was made by a woman." Is there a fundamental logical difference between the two statements as reformulated?

SUGGESTED READINGS

FRIEDMAN, MICHAEL (1974), "Explanation and Scientific Understanding," *Journal of Philosophy 71*: 5–19. Reprinted in Pitt (1988). This is the original statement of the unification approach to scientific explanation. Although brief highly technical parts appear near the end, the article contains interesting general discussion of the basic issues.

HEMPEL, CARL G. (1942), "The Function of General Laws in History," *Journal of Philosophy 39*: 35–48. Reprinted in Hempel (1965b). Hempel's original essay on explanation in history.

———. (1959), "The Logic of Functional Analysis," in Llewellyn Gross (ed.), *Symposium on Sociological Theory*. New York: Harper & Row, pp. 271–307. Reprinted in Hempel (1965b). Hempel's original essay on functional explanation.

———. (1962), "Explanation in Science and in History," in Robert G. Colodny (ed.), *Frontiers of Science and Philosophy*. Pittsburgh: University of Pittsburgh Press, pp. 7–33. A lucid and highly readable brief presentation of Hempel's basic views.

———. (1965a), "Aspects of Scientific Explanation," in Hempel (1965b), pp. 331–496. Hempel's magisterial comprehensive monograph on scientific explanation.

———. (1965b), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: The Free Press. Part 4 contains four classic articles on scientific explanation.

———. (1966), *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall. Chapters 5–8 offer an extremely elementary and highly readable introduction to Hempel's views.

HEMPEL, CARL G. and PAUL OPPENHEIM (1948), "Studies in the Logic of Explanation," *Philosophy of Science 15:* 135–175. Reprinted in Hempel (1965b) and in Pitt (1988). This is *the* modern classic on scientific explanation. Parts 1 and 3 are especially important.

HUMPHREYS, PAUL (1981), "Aleatory Explanation," *Synthese 48:* 225–232. An introductory account of a new approach to statistical explanation.

JEFFREY, RICHARD C. (1969), "Statistical Explanation vs. Statistical Inference," in Nicholas Rescher (ed.), *Essays in Honor of Carl G. Hempel.* Dordrecht: Reidel, pp. 104–113. Reprinted in Wesley C. Salmon and others (1970) *Statistical Explanation and Statistical Relevance*. Pittsburgh: University of Pittsburgh Press. This philosophical gem explicitly raises the question of whether explanations are arguments.

PITT, JOSEPH C. (ed.) (1988), *Theories of Explanation*. New York: Oxford University Press. An anthology containing a number of important articles.

RAILTON, PETER (1981), "Probability, Explanation, and Information," *Synthese 48:* 233–256. A clear and careful elementary exposition of a mechanistic approach to probabilistic explanation.

SALMON, WESLEY C. (1978), "Why Ask, 'Why'?—An Inquiry Concerning Scientific Explanation," *Proceedings and Addresses of the American Philosophical Association 51:* 683–705. An introductory discussion of explanation by means of unobservable entities.

———. (1982), "Comets, Pollen, and Dreams: Some Reflections on Scientific Explanation," in Robert McLaughlin (ed.), *What? Where? When? Why?* Dordrecht: Reidel, pp. 155–178. A popularized discussion of some of the basic issues concerning scientific explanation.

———. (1990), *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press. A historical survey of philosophical developments since the classic Hempel-Oppenheim (1948) essay.

SCRIVEN, MICHAEL (1959), "Explanation and Prediction in Evolutionary Theory," *Science 130:* 477–482. One of the first strong challenges to Hempel's account.

———. (1962), "Explanations, Predictions, and Laws," in Herbert Feigl and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*. Vol. 3, *Scientific Explanation, Space and Time*. Minneapolis: University of Minnesota Press, pp. 170–230. Another strong attack on the received view of scientific explanation.

VAN FRAASSEN, BAS C. (1980), *The Scientific Image*. Oxford: Clarendon Press. Chapter 5 presents an influential treatment of the pragmatics of explanation.

WRIGHT, LARRY (1976), *Teleological Explanations*. Berkeley and Los Angeles: University of California Press. An important treatment of teleological and functional explanation.