

The Fourth Decade (1978–87)

A Time of Maturation

The history of the last decade is the hardest to recount, especially for one who has been deeply involved in the discussions and controversies. It is difficult to achieve perspective at such close range. Nevertheless, I think certain features are discernible. It is a period during which several different lines of thought achieved relatively high degrees of maturity. For example: (1) The role of causality in scientific explanation has been pursued in far greater detail than previously. (2) Views on the nature of statistical explanation and its relationship to causal explanation have become much more sophisticated. (3) Our understanding of the appeal to unobservables for purposes of explanation has been considerably advanced. (4) The pragmatics of explanation—which posed fundamental points of controversy from the beginning of the second decade—has been investigated more deeply and with more precision than ever before. (5) The question of the relationship between descriptive knowledge and explanatory knowledge has been examined more closely. (6) Perhaps most important, the explicandum has received significant and much-needed clarification during the decade just passed. Let us begin with this last item.

4.1 New Foundations

During the third decade of our chronicle—the one entitled “Deepening Differences”—there was an increasing awareness of the difficulties in the received view. As the hegemony crumbled philosophers looked more closely at the foundations. I recall vividly my own feeling that, in discussions of scientific explanation, the explicandum was in serious need of clarification. Too often, I felt, those who wrote on the subject would begin with a couple of examples and, assuming that these particular cases made the concept clear enough, would proceed to the task of explicating it. Referring to the exemplary job Carnap had done on clarification of the explicandum in *Logical Foundations of Probability* (chaps. I, II, and IV), I declared that the same sort of thing needed to be done for scientific explanation.

I was by no means the only philosopher to feel an urgent need for clarification of the explicandum in the mid-1970s. Michael Friedman's "Explanation and Scientific Understanding" (1974) was a seminal contribution. So also were two articles published near the close of the third decade—one by D. H. Mellor (1976), the other by Alberto Coffa (1977). Largely as a result of these latter two articles, I became acutely aware of the need to distinguish three fundamentally distinct conceptions of scientific explanation—modal, epistemic, and ontic (first presented in W. Salmon 1982).

Mellor adopts a *modal conception* of scientific explanation. Suppose some event E occurs; for all we know at the moment, it might or might not have happened. We explain it by showing that, given other circumstances, it *had to happen*. The modal conception has been advocated by a number of philosophers, but almost always in a deterministic context. If determinism is true, according to most who embrace this view, then all events are in principle explainable. If indeterminism is true, some events will not be amenable to explanation, even in principle; only those events that are necessitated by preceding conditions can be explained. In either case, however, it is not sufficient to show merely *that* the event-to-be-explained had to happen. By citing antecedent circumstances and universal laws one shows *why it had to happen*—by virtue of what it was necessitated.

But Mellor puts a novel twist on the modal conception. He claims that there can be explanations of events that are irreducibly probabilistic. By showing that an event has a high probability relative to preceding circumstances, Mellor claims, we close the gap to some extent, so to speak. There are degrees of necessitation. If an event is completely necessitated, its occurrence is fully entailed by laws and explanatory facts. If it is not fully necessitated, its occurrence is partially entailed by explanatory facts. The greater the degree of partial entailment, the better the explanation.

The concept of *partial entailment* is intuitively appealing to many philosophers who want to follow the Carnapian strategy of constructing an inductive logic on a logical interpretation of probability. The idea is to build inductive logic on the relation of partial entailment in much the same way as deductive logic can be built on the relation of full entailment. Although Carnap never used this concept in his precise explications, he does make informal reference to it (1950, 297). This concept strikes me as one of dubious value. In order to have a measure of degree of partial entailment, one is required to select an a priori measure that is tantamount to assigning prior probabilities to all of the statements that can be formulated in the language to which the inductive logic is to be applied. Given the wide range of choices for such a measure—a nondenumerable infinity (Carnap 1952)—the a priori character of this choice makes it egregiously arbitrary.¹ Thus, it seems to me, partial entailment cannot be construed as degree of necessitation in any way that is useful to the modal conception of scientific explanation. The modal concep-

tion therefore appears to require the domain of legitimate scientific explanations to be restricted to deductive explanations. Whether this restriction is tolerable in an era in which physics gives strong indication that we live in an indeterministic universe is an extremely serious question. We shall have to consider this issue in greater detail in §4.9 on deductivism. In my view it is an untenable position.

In the year following the appearance of Mellor's paper, Coffa offered a perceptive contrast between the I-S (inductive-statistical) and the S-R (statistical-relevance) models of explanation (1977).² On Hempel's account, we recall, an explanation—deductive or inductive—is an argument. In an I-S explanation, the relation of explanans to explanandum is an inductive or epistemic probability. In a good I-S explanation that probability is high, the higher the better. Explanations—both deductive and inductive—show that the event to be explained *was to be expected*. In view of these considerations, Coffa described Hempel's conception as *epistemic*. The S-R model is conceived in terms of objective probabilities; my formulation was given in terms of relative frequencies. In his closely related dispositional theory of inductive explanation Coffa appealed to propensities.³ On either account, high probability per se has no particular virtue; what matters is to get the objective probabilities right. Consequently, Coffa characterized the conception underlying these models as *ontic*.

The first year of the fourth decade saw my first serious effort in print at *clarification of the explicandum*. It was embodied in "Why Ask, 'Why?'—An Inquiry Concerning Scientific Explanation" (1978), my Presidential Address to the Pacific Division of the American Philosophical Association. Attention was focused on two basic concepts of explanation, the inferential (à la Hempel) and the causal (à la Scriven). I tried to exhibit the powerful intuitions that underlie each of them. The most appealing examples from the standpoint of the former are those in which one or more regularities are explained by derivation from more comprehensive regularities—the Newtonian explanation of Kepler's laws, or the Maxwellian explanation of the laws of optics on the basis of electromagnetic theory. Yet the vast majority of examples given by Hempel and other supporters of this conception are explanations of particular facts. Moreover—a point I did not make but should have emphasized—neither the Hempel-Oppenheim article nor Hempel's "Aspects" essay even attempts to offer an account of explanations of laws.

The most persuasive examples from the standpoint of the causal conception are explanations of particular occurrences, often in the context of practical applications. Explanations of airplane crashes and other disasters provide a wealth of instances. The main shortcoming of the causal conception, I argued, was its lack of any adequate analysis of causality. I tried to provide the foundations of such an account in terms of *causal processes*, *conjunctive forks*, and *interactive forks*. It was a theme to which I have returned a number of times during the last ten years.

Coffa's distinction between the epistemic and ontic conceptions of explanation

arose in the context of statistical explanation. In "Why Ask, 'Why?'" I sought to apply it more generally to the distinction between the inferential conception and the causal conception. As it seemed to me, the epistemic conception is oriented toward the notion of scientific expectability, while the ontic conception focuses upon the fitting of events into natural regularities. Those regularities are sometimes, if not always, causal. Thinking in terms of this distinction between the epistemic and ontic conceptions, I was surprised to find that, in "Aspects of Scientific Explanation," Hempel offers brief characterizations of explanation in general near the beginning and at the end. The two do not agree with each other.

In his initial discussion of D-N (deductive-nomological) explanation, Hempel says that an explanation of this type "may be regarded as an argument to the effect that the phenomenon to be explained . . . was to be expected in virtue of certain explanatory facts" (1965, 336). A bit later this conception is applied to statistical explanation as well. This characterization obviously reflects the epistemic conception. At the conclusion of this essay he sums up his theory of scientific explanation in these terms: "The central theme of this essay has been, briefly, that all scientific explanation involves, explicitly or by implication, a subsumption of its subject matter under general regularities; it seeks to provide a systematic understanding of empirical phenomena by showing that they fit into a nomic nexus" (1965, 488). For the sake of causal explanation, I would be inclined to rephrase the general characterization slightly: it seeks to provide a systematic understanding of empirical phenomena by showing *how* they fit into a *causal* nexus. Nevertheless, either way, the ontic conception is being expressed. Hempel did not, I suspect, notice the differences between his initial and final formulations.

In his doctoral dissertation—a work that is, in my opinion, quite possibly the best thing written on scientific explanation since Hempel's "Aspects" essay—Peter Railton (1980) makes an observation regarding Hempel's theory that is closely related to Coffa's distinction between epistemic and ontic conceptions. Approaching the received view in an extraordinarily sensitive way, Railton shows that Hempel's thesis to the effect that explanations explain by conferring *nomiic expectability* on the explanandum cannot be maintained. The problem is that nomic expectability involves two components, *nomicity* and *expectability*, that can conflict with each other. A particular event, such as a spontaneous radioactive decay, may be rather improbable, yet we know the ineluctably statistical laws that govern its occurrence. The nomic side is fulfilled, but the expectability side is not. Hempel chose to reject, as nonexplanatory, any account that renders the event-to-be-explained improbable.⁴ Railton argues that it would be better to accept such accounts as explanatory, provided they fulfill certain general conditions, and to take *nomicity* rather than *expectability* as the key to scientific explanation. If one takes that tack—as both Coffa and Railton clearly realized—it amounts to relinquishing the epistemic conception in favor of the ontic.

According to the ontic conception, the events we attempt to explain occur in

a world full of regularities that are causal or lawful or both. These regularities may be deterministic or irreducibly statistical. In any case, the explanation of events consists in fitting them into the patterns that exist in the objective world. When we seek an explanation, it is because we have not discerned some feature of the regular patterns; we explain by providing information about these patterns that reveals how the explanandum-events fit in. Along with Coffa, both Railton and I endorse the ontic conception. We all maintain that explanations reveal the mechanisms, causal or other, that produce the facts we are trying to explain. The greatest difference between Railton and me concerns the degree to which explanations must be causal. His view is more lenient than mine with regard to noncausal explanation.

During the latter part of 1978 I had the great privilege of visiting Australia and offering a seminar on scientific explanation at the University of Melbourne in the Department of History and Philosophy of Science (the oldest such department in the world, I believe). This visit afforded the opportunity to think through more fully these foundational questions and to discuss them with a number of colleagues and students at Melbourne and at several other Australian universities as well. During this visit to Australia I composed the first drafts of several chapters of *Scientific Explanation and the Causal Structure of the World* (1984).

Taking the cues provided by Coffa and Mellor, it seemed to me that we could distinguish three basic conceptions of scientific explanation—modal, epistemic, and ontic—that could be discerned in Aristotle, and that have persisted down through the ages. In the context of Laplacian determinism they seem to merge harmoniously; in the context of statistical explanation they diverge dramatically. The modal conception (I claim, *pace* Mellor) precludes statistical explanation—except, perhaps, as some sort of incomplete explanation. The epistemic conception, as it occurs within the received view, requires high inductive probabilities. The ontic conception demands objective probabilities, whether high, middling, or low.⁵ I offered a brief general discussion of these differing conceptions in “Comets, Pollen, and Dreams: Some Reflections on Scientific Explanation,” a rather popularized account that was published in a collection of essays that grew out of my trip to Australia (Salmon 1982).

The foregoing tripartite division was a fairly serviceable crude classification scheme for theories of scientific explanation, but it required refinement. In the far more detailed treatment of the three basic conceptions (1984, chap. 4), I pointed out that the epistemic conception has three different versions—inferential, information-theoretic, and erotetic. Because of his insistence that explanations are arguments, it is appropriate to dub Hempel's conception the *inferential version of the epistemic conception*. This version represents the received view of scientific explanation.

Early in the third decade, James G. Greeno (1970) and Joseph Hanna (1969) began offering information-theoretic accounts of scientific explanation. Greeno

took the notion of *transmitted information* as a basis for evaluating the explanatory adequacy of statistical theories. Given the fact that information is the key concept, he clearly is adopting an epistemic approach; an appropriate designation would be the *information-theoretic version* of the *epistemic conception*. A crucial difference between his model and Hempel's I-S model is that information transmitted reflects a relevance relation. Another major difference is that Greeno evaluates the explanatory power of laws or theories, but he does not provide a method for evaluating particular explanations of particular facts. This global feature constitutes an affinity between Greeno's information-theoretic account and Friedman's unification account. It seems to me that, if one wants to maintain an epistemic conception, even in the face of serious criticisms of Hempel's models, Greeno's approach is the most promising. Kenneth Sayre (1977) and Joseph Hanna (1978, 1981, 1983) have made subsequent important contributions to it.

Still another version—the *erotetic version*—of the epistemic conception must be distinguished. The term “erotetic” was chosen because the logic of questions has been known traditionally as erotetic logic. This version was suggested by Braithwaite when he remarked that “an explanation, as I understand the use of the word, is an answer to a ‘Why?’ question which gives some intellectual satisfaction” (1953, 348–49), but he does not develop this approach. Bromberger's work on why-questions (1966), in contrast, involves a sustained effort to elaborate the nature of why-questions and their relations to scientific explanation. But the best-known articulation of this version of the epistemic conception can be found in Bas van Fraassen's provocative work *The Scientific Image* (1980, chap. 5). This approach has obvious connections with Bromberger's earlier work, but a few major differences should be noted. First, although Bromberger presented a theory of explanations as answers to why-questions, he did not consider it a comprehensive treatment, for he flatly denied that all requests for explanations can be phrased as why-questions. Along with a number of other philosophers, Bromberger claims that some explanations are answers to (among others) how-possibly-questions, and that these are different from explanations that are answers to why-questions. Since erotetic logic deals with all kinds of questions, that fact does not disqualify Bromberger as a representative of the erotetic version of the epistemic approach. Van Fraassen, in contrast, affirms the view that all explanations can be considered answers to why-questions. Second, Bromberger made no attempt to deal with statistical explanations; van Fraassen's theory is intended to include explanations of this type. We shall discuss van Fraassen's theory in detail in §4.4.

4.2 Theoretical Explanation

Another major topic of “Why Ask, ‘Why?’?” involved the appeal to unobservable entities for purposes of scientific explanation. The issue was by no means new. Within the hegemony of logical empiricism there was what might well be

termed a "received view of theories." The main idea was that at the most basic level we have the particular empirical facts revealed by observation, at the next level are empirical generalizations concerning observables, and at the next higher level are theories that seem to make reference to unobservable entities. According to the received view of scientific explanation, the empirical laws explain the observed phenomena, and the theories explain the empirical laws. For example, various observed facts about pressures, temperatures, and volumes of gases are explained by the empirical ideal gas *law*, and that law is explained by the molecular-kinetic *theory*. There may, of course, be still higher level theories that explain the lower level theories. Wilfrid Sellars disparagingly dubbed this account of observable facts, empirical laws, theories, and the explanatory relationships among them the "layer cake" account. It was spelled out rather explicitly in the first four chapters of Braithwaite's *Scientific Explanation*.

During the nineteenth century there had been a good deal of resistance on the part of many scientists and philosophers to the notion that such microentities as atoms or molecules actually exist, or, at any rate, to the notion that we could possibly know anything about them if they do exist. Even those scientists who recognized the utility of the molecular-kinetic theory sometimes regarded such entities merely as useful fictions. This viewpoint is known as *instrumentalism*. The theory is a useful instrument for making scientific predictions (see Gardner 1979).

In the early part of the twentieth century these concerns about our ability to have knowledge of unobservable entities were transformed into questions about the meaningfulness of theories. Operationists and logical positivists denied that utterances putatively about unobservable objects could be scientifically meaningful. Such logical empiricists as Carnap and Hempel made serious efforts to show how—without abandoning empiricist principles—scientific theories could be meaningfully construed. An excellent account of these developments can be found in Hempel's classic essay, "The Theoretician's Dilemma" (1958). At the end of this paper he argues that theories are required for "inductive systematization," but I am not sure that he can establish more than the heuristic value of theories. Even the instrumentalist can cheerfully admit that theories are extremely useful in science. The basic question is whether the unobservable entities to which they seem to refer actually exist.

It is imperative, I believe, to separate the question of the existence of entities not directly observable by means of the unaided human senses from the issue of the meaningfulness of a theoretical vocabulary. Logical empiricists like Carnap and Hempel suggested that the terms of our scientific language can be subdivided into two parts—an observational vocabulary, containing such terms as "table," "dog," "red," "larger than," etc., and a theoretical vocabulary containing such terms as "electron," "atom," "molecule," "gene," "excited state (of an atom)," etc. The viability of a sharp observational-theoretical distinction was frequently called into question, but that particular problem need not detain us now. The instrumen-

talism issue can be formulated without reference to any such distinction in the scientific vocabulary.

Consider, for example, the epoch-making work of Jean Perrin on Brownian movement in the first decade of the twentieth century. To conduct his experiments he created tiny spheres of gamboge (a bright yellow resinous substance) less than a micrometer in diameter, and he accounted for their motions in terms of collisions with even smaller particles. These experiments were, as we shall see, crucial to Perrin's argument regarding the reality of molecules. Notice that I have formulated the key statements about unobservable entities without going beyond the observational vocabulary.⁶

By the time the statistical-relevance model of scientific explanation had been fairly completely articulated (circa 1970) I was aware of the fact that it was not obviously capable of accommodating theoretical explanation. At that time, James Greeno, who had developed an information-theoretic approach to statistical explanation, presented a paper (1971) in a Philosophy of Science Association symposium in which he tried to show how appeal to theories could yield an increase in information. In my comments in that symposium (1971) I showed that his approach, attractive as it was, would not work. I then set about trying to provide one that would.

In 1973 a conference on explanation (not just scientific explanation) was held at the University of Bristol; the proceedings were published in (Körner 1975). Upon receiving an invitation, but (regrettably) before writing the paper, I proposed the title "Theoretical Explanation" for my contribution. In the end I found I had written a paper on causal explanation, in which I discussed at some length appeals to continuous causal processes and common cause arguments, but which failed to yield any solid result about unobservables. In "Why Ask, 'Why?'" I thought I had the fundamentals of an approach that would work.

Unconvinced by the various arguments about theoretical realism that had been offered by philosophers from the 1930s to the 1970s, I undertook to find out what considerations convinced natural scientists of the existence of such unobservables as atoms and molecules. Without having been aware at that time of the historical importance of Avogadro's number N , I did recognize that it provided a crucial link between the macrocosm and the microcosm. With the help of N , one could calculate micro-quantities from macro-quantities and conversely. From the mass of a mole of any given substance, for example, N gives us immediately the mass of a molecule of that substance. Thus, it seemed to me, the ascertainment of N was a good place to start.

A first and most obvious way to get at the value of N is through the phenomenon of Brownian movement. According to the molecular-kinetic theory of gases, as Einstein and Smoluchowski showed in 1905–06, the motion of a Brownian particle suspended in a gas is the result of random bombardment by the molecules of the gas. Assuming that the gas and the Brownian particles are in thermal

equilibrium, it follows that the average kinetic energy of the Brownian particles is equal to the average kinetic energy of the molecules. As Perrin remarked at about the same time, the dance of the Brownian particles represents qualitatively the random motion of the molecules. By ascertaining the mass of the Brownian particle, the average velocity of the Brownian particle, and the average velocity of the molecules, one can compute the mass of the molecule directly. In practice the situation is a bit more complicated. Although the mass of the Brownian particle and the average velocities of the molecules are quite directly measurable, the average velocity of the Brownian particle is not, for it changes its direction of motion too rapidly. But by indirect means—basically, observation of rates of diffusion of Brownian particles—what amounts to the same method can be applied. As a result we know the mass of a molecule of a gas and Avogadro's number N , the number of molecules in a mole of that gas. This type of experiment was done in the early years of the twentieth century by Perrin.⁷

The instrumentalist can easily reply to the preceding consideration by pointing out that since the advent of the molecular-kinetic theory we have known that it is useful to think of a gas as composed of little particles that move at high speeds and collide with one another and with the walls of the container. Now, it can be added, the Brownian particle behaves as if it is being bombarded by these little particles, and, indeed, we can say that the gas behaves as if it is composed of a certain number of these tiny particles. But all of this does not prove the reality of molecules; it shows that the molecular kinetic theory is an excellent instrument for predicting the behavior of gases (and of the Brownian particles suspended in them). Thus, the manifest success of the molecular-kinetic theory did not constitute compelling evidence for the existence of molecules. As a matter of historical fact many serious and knowledgeable physical scientists at the turn of the century did not believe that atoms and molecules are real.

The way out of this difficulty lies in the fact that N can be ascertained in a variety of ways. In "Why Ask, 'Why?'" I mentioned a determination by means of electrolysis as an example of a totally different experimental approach. If an electric current passes through a solution containing a silver salt, an amount of metallic silver proportional to the amount of electric charge passing through the solution is deposited on the cathode. The amount required to deposit one mole of a monovalent metal (such as silver) is known as a faraday. A faraday is found by experiment to be 96,487 coulombs. If that charge is divided by the charge on the electron, empirically determined by J. J. Thomson and Robert Millikan to be -1.602×10^{-19} coulombs, the result is N . A faraday is simply Avogadro's number of electron charges. Superficially, the phenomena involved in these two experiments are entirely different, yet they agree in the numerical value they yield for N . I suggested that this agreement, within experimental error, of the value derived via the study of Brownian movement with the value derived via the electrolysis experiment suggests that the particulate character of matter—the reality

of such things as molecules, atoms, ions, and electrons—is a common cause of this agreement. I realized at the time that there are many other ways of ascertaining N, but mentioned just two in order to illustrate the common cause character of the argument.

At some time during 1978 I became aware of Mary Jo Nye's superb historical account of the work of Jean Perrin (1972) and, through that, of Perrin's own semi-popular treatment (1913, English translation 1923). In the period between 1905 and 1913 Perrin had done a spectacular set of experiments on Brownian movement and the determination of N. Combining his findings with those of other workers, Perrin presents a table, near the close of his book, listing the results of *thirteen distinct methods* for ascertaining N, and notes the striking agreement among them. Immediately after the table he remarks:

Our wonder is aroused at the very remarkable agreement found between values derived from the consideration of such widely different phenomena. Seeing that not only is the same magnitude obtained by each method when the conditions under which it is applied are varied as much as possible, but that the numbers thus established also agree among themselves, without discrepancy, for all the methods employed, the real existence of the molecule is given a probability bordering on certainty. (1923, 215–16)

This was, indeed, the argument that convinced virtually every physical scientist of the existence of unobservable entities of this sort.

In *Scientific Explanation and the Causal Structure of the World* (213–27) I tried to spell out this common cause argument in some detail. However, instead of dealing with thirteen different ways of ascertaining N, I confined my attention to five: Brownian movement, electrolysis, alpha radiation and helium production, X ray diffraction by crystals, and blackbody radiation. They constitute a highly diverse group of experiments. I thought it fitting, given St. Thomas Aquinas's five ways of demonstrating the existence of God, that I should cite the same number to establish the existence of atoms and molecules. I consider the argument that convinced the scientists in the early part of the twentieth century philosophically compelling.

4.3 Descriptive vs. Explanatory Knowledge

What good are explanations? This question has been asked—explicitly or implicitly—on innumerable occasions over the years, and it has received a variety of answers. During the first half of the twentieth century, scientific philosophers were concerned to refute such answers as, “Explanations inform us of the ultimate purposes for which the world was created,” “Explanations reveal the underlying essences of the things we find in the world,” or “Explanations exhibit the vital forces or entelechies within living beings.” As I remarked near the beginning, the

heavy theological and metaphysical involvements of philosophy during that period led some scientists and philosophers to reject altogether the notion that science has anything to do with explanation. Those who did not want to relinquish the claim that science can furnish legitimate explanations were at pains to make it clear that explanatory knowledge is part of our empirically based descriptive knowledge of the natural world. The classic paper by Hempel and Oppenheim, as well as Hempel's "Aspects" paper, were efforts to delineate exactly what sort of descriptive knowledge constitutes explanatory knowledge. These essays were designed to show that legitimate scientific explanations could be had without appealing to superempirical facts or agencies. This is what the received view was all about.

At the same time, these treatments of explanation rejected such psychologistic answers as "Explanations increase our understanding by reducing the unfamiliar to the familiar," or "Explanations ease our intellectual discomforts and make us feel more at home in the world" (see H-O 1948, §4; W. Salmon 1984, 12-15). Great caution must be exercised when we say that scientific explanations have value in that they enable us to understand our world, for *understanding* is an extremely vague concept. Moreover—because of the strong connotations of human empathy the word "understanding" carries—this line can easily lead to anthropomorphism. The received view was concerned also to show that scientific explanations can be given without indulging in that intellectual vice either.

These considerations lead, however, to a deeply perplexing puzzle. If explanatory knowledge does not exceed the bounds of descriptive knowledge, in what *does* it consist? What do we have when we have explanations that we did not already have by virtue of our descriptive knowledge? What is the nature of the understanding scientific explanations are supposed to convey? The full force of this problem did not strike me until about 1978, and I addressed it explicitly in "Why Ask, 'Why?'" (1978). At about the same time, van Fraassen was asking the same question. We shall consider his answer, which is quite different from mine, below.

The way I posed the problem is this. Suppose you were Laplace's demon, possessing a complete description of the world at one particular moment, knowing all of the laws, and having the ability to solve the mathematical problems required to predict or postdict everything that ever has happened or ever will happen. Such knowledge would appear to be descriptively complete; what else would be involved in having explanations?

I now think this was a particularly inept way to put the question, for one obvious answer is that the demon would have no occasion to ask "Why?" Van Fraassen would, I suspect, agree with this response. As our discussion of Bromberger's work has revealed, one asks a why-question only as a result of some sort of perplexity—recall his p- and b-predicaments. The demon would not be in any such predicament. So we should drop the fantasy of the Laplacian demon. Why-

questions are raised only in contexts in which our knowledge is incomplete. Nevertheless, the fundamental question remains, and remains important. What sort of information is explanatory information? How, if at all, does explanatory knowledge differ from other types of descriptive knowledge? The three basic conceptions discussed in §4.1 offer distinct answers to this question.

For the modal conception a straightforward answer is available. Explanatory knowledge adds a modal dimension to our descriptive and predictive knowledge. Explanatory knowledge is knowledge of what is necessary and what is impossible.⁸

In "Why Ask, 'Why?'" I attempted to respond on behalf of the ontic conception. Even though the question was badly put, the answer points, I think, in the right direction. I suggested that, in addition to purely descriptive knowledge, one would need causal knowledge: recognition of the difference between causal and noncausal laws, the difference between causal processes and pseudo-processes, and the difference between causal interactions and mere spatio-temporal coincidences. At present I would be inclined to phrase the answer somewhat differently in terms of laying bare the underlying mechanisms, but the basic idea seems to me sound. According to the ontic conception, explanatory knowledge is knowledge of the causal mechanisms, and mechanisms of other types perhaps, that produce the phenomena with which we are concerned.

Since it comes in three distinct versions, the epistemic conception requires three answers. According to the received view, explanations involve the subsumption of the fact to be explained under some kind of lawful regularity, universal or statistical. In that way they provide *nomic expectability*. But, one might ask, do we not already have nomic expectability as part of our descriptive and predictive knowledge?⁹ The received view seeks assiduously to avoid any incursion into metaphysics or theology for purposes of scientific explanation, but it is hard to see just what constitutes the explanatory import of an explanation on that conception. Indeed, given the original strong version of Hempel's thesis of the symmetry between explanation and prediction, this point obviously holds. The *only* difference between explanation and prediction is pragmatic. If we know that the fact described in the conclusion of an argument conforming to one of the models of explanation actually obtains, the argument constitutes an explanation of that fact. If we do not already know that this fact obtains, the very same argument constitutes a prediction.

As Israel Scheffler pointed out around the beginning of the second decade, it is highly implausible to maintain that every legitimate scientific prediction can serve, under suitable pragmatic conditions, as an explanation. One major reason, as we saw, is that inductive arguments from particulars to particulars—not involving any law—can provide reliable predictions. The symmetry thesis must therefore be amended, at the very least, to the claim that predictions based upon laws can function as explanations in certain epistemic contexts. Thus, the appeal

to *nomic regularities* is the crucial feature of explanation. Since it seems reasonable *prima facie* to claim that law statements constitute a crucial part of our descriptive knowledge, we still must ask in what way explanations, as characterized by the received view, have *explanatory import*.

It is useful to recall, in this connection, the fundamental point made by Railton concerning nomic expectability—namely, sometimes nomicity and expectability conflict. If one goes with expectability, thus fully preserving the epistemic character of the conception, descriptive and explanatory knowledge seem indistinguishable. If one opts for nomicity instead, one ends up—as Hempel did at the conclusion of “Aspects”—in the ontic conception.

Before going on to consider the other two versions of the epistemic conception, let us pause to compare the three responses already given. When we appeal to modality—to necessity or impossibility—it may involve either of two distinct approaches. First, one might look upon the modalities as metaphysical categories, residing in a realm separate from the domain that can be empirically investigated. This construal would exile explanation from science. Second, one might say that nothing beyond physical necessity and impossibility is involved, and that these come directly from the laws of nature. If law-statements constitute part of our descriptive knowledge, then the modal conception makes no appeal to anything beyond our descriptive knowledge. Notice that, on this interpretation, the modal conception and the received view are in complete agreement on the status of D-N explanation. The received view differs from the modal conception only insofar as it admits some form of statistical explanation.

When we look at the inferential version of the epistemic conception—involving *nomic* expectability—we see immediately that the status of laws becomes crucial. Consequently, we must ask, what is the difference between true lawlike generalizations (of either the universal or statistical variety) and true accidental generalizations? Recall our previous example comparing statements about massive spheres of gold with statements about massive spheres of uranium. Is there any objective difference between lawful and accidental generalizations, or is it merely a matter of our greater confidence in one as opposed to the other? Alternatively, do we want to characterize laws in terms of relations among universals? If so, this takes the distinction between laws and nonlaws, and, consequently, the characterization of explanation, out of the realm of empirical science. It appears, therefore, that the choice, for the adherent of the received view, is between a heavily pragmatic theory (such as we found in Rescher) and an extrascientific metaphysical one. Neither choice, I believe, captures the intent of the received view.

According to the ontic conception, there is a further gap between explanation and prediction. As we noted much earlier in connection with the famous barometer example, the sharply falling barometric reading is a satisfactory basis for predicting a storm, but contributes in no way to the explanation of the storm. The

reason is, of course, the lack of a direct causal connection. For the ontic conception, therefore, mere subsumption under a law is not sufficient for explanation. There must be, in addition, a suitable causal relation between the explanans and the explanandum—at least as long as we steer clear of quantum mechanical phenomena.

The basic question, for this conception, is the status of causality. If we construe causal relations in extrascientific metaphysical terms we will banish explanation from science. If we follow a purely Humean tack, construing causality strictly in terms of constant conjunction (see Mackie 1974),¹⁰ we will make the ontic approach identical with the received view.¹¹ The approach I adopted in “Why Ask, ‘Why?’?” involved an appeal to causal processes and causal interactions (see §3.6 above). Causal processes are distinguished from pseudo-processes in terms of the ability to transmit marks. I attempted to give an entirely empirical construal of the causal concept of transmission in terms of the *at-at* theory. I distinguished causal interactions from mere spatio-temporal intersections of processes in terms of mutual modification of the processes involved. To my great regret, I found no way of carrying out these explications without the use of counterfactual conditionals (W. Salmon 1984, 148–50, 171–74). I have no analysis of counterfactuals, though I do offer a method for testing the truth of the kinds of counterfactuals that are invoked in this context. I am not terribly dissatisfied with this state of affairs, but doubtless other philosophers will not be as easily satisfied on this score.

We have known for a long time that three sets of issues are tightly intertwined: modalities, laws, and counterfactuals (see W. Salmon, 1976). It seemed that any two of the three could be satisfactorily analyzed in terms of the third, but it is difficult to produce a satisfying analysis of any one of them that did not invoke at least one of the other two. Circularity is, consequently, always a serious threat. The most popular current approach to counterfactuals seems to be one that appeals to possible worlds (see Lewis 1973).¹² It has two major shortcomings in my opinion. First, the postulation of the existence of myriad possible worlds, distinct from our actual world, takes us deep into the superempirical.¹³ Second, evaluation of the similarity of possible worlds—which is essential to the analysis of counterfactuals—requires an appeal to laws. So, it has not broken us out of the circle. It is quite interesting to note, then, that (1) the modal conception attaches itself to modality, (2) the epistemic conception, inferential version, depends upon laws, and (3) the ontic conception—as far as I can see, at least—appeals to counterfactuals. Each conception has its cross to bear. Yet in that cross may lie the ‘something extra’ beyond sheer descriptive knowledge that is supposed to provide scientific understanding.

Let us now return to the remaining two versions of the epistemic conception. The information-theoretic version faces almost the same difficulty as did the inferential version in explicating the distinction between purely descriptive knowl-

edge and explanatory knowledge. According to the information-theoretic approach the explanatory value of a law or theory is measured in terms of its efficacy in improving predictive power. That is just what information transmitted amounts to. There is, nevertheless, a possible approach that can be shared by the inferential version and the information-theoretic version. In discussing explanation, prediction, the status of theories, and other related concepts, Hempel has often referred generally to *systematization*—both *deductive systematization* and *inductive systematization*. This suggests the possibility of construing explanatory force, not in terms of some extra type of knowledge, but rather, in terms of the organization of the descriptive knowledge we have or can procure.

This way of looking at explanatory force bears a striking resemblance to Friedman's unification theory of explanation. According to that view, explanations improve our understanding through the unification of our knowledge. Our understanding is increased when we can reduce the number of independent assumptions we have to make about the world. Another way to look at unification is in terms of information theory. When a great deal of information about the world is contained in a short message, we have increased understanding. Either way, what is crucial to explanation is not some particular kind of explanatory knowledge, but, rather, the way in which our descriptive knowledge is organized.

According to the erotetic version of the epistemic conception, as developed by van Fraassen, there is no fundamental difference between descriptive knowledge and explanatory knowledge. On van Fraassen's theory an explanation is simply an *answer* to a why-question; it is nothing other than descriptive information that, *in a given context*, answers a particular type of question. Whether a piece of information constitutes explanatory knowledge depends solely upon the context in which it is furnished. Thus, whatever distinction there is between descriptive and explanatory knowledge is entirely pragmatic.

When we use the term "descriptive knowledge," there is serious danger of equivocation. In one meaning of the term, describing the world or any part of it consists only in reporting what is apparent to the senses. Thus, one could describe a given volume of air as warm or cold, calm or windy, clear or hazy, etc. One might also include such readily detectable features as moisture content and pressure. Such a description would *not* include the fact that air consists of molecules of various gases such as nitrogen, oxygen, and carbon dioxide, or that these molecules are made up of atoms having certain characteristics. A complete description of this sort would be a description solely of appearances. In another sense, a complete description of an entity would include all kinds of facts about it, directly observable or not. Such a description of the same body of air, if complete, would specify how many molecules there are of each type, the atomic constitution of molecules of each type, and facts about the collisions of molecules. This is the sort of description Laplace's demon would have possessed. When we raise the question about the relationship between descriptive knowledge and ex-

planatory knowledge, we must be careful to indicate in which of the foregoing two senses we are construing the word "descriptive." This has a crucial bearing upon the theory of scientific explanation.

In the first chapter of his doctoral dissertation, Coffa discussed in detail the nature of explication, and he devoted serious attention to the clarification of the explicandum he was trying to explicate. He called attention to

a centuries-old philosophical tradition, sometimes referred to by the name of 'instrumentalism', that has denied the claim that science has explanatory power. For instrumentalists there are no scientific explanations. Science is acknowledged to have a number of virtues, but none of them is associated with the production of a better understanding of what goes on in the world. Science could be a source of predictive power, the foundation of technology, an unlimited fountainhead of aesthetic pleasure, the maximal organizational principle of experience; but it would not be a source of understanding. For the instrumentalist science is a prediction machine with pleasant, but only psychological side-effects.

The notion of explanation we want to explicate is one whose correct applicability to a scientific argument is denied by instrumentalists. An alleged theory of explanation (i.e., an explication of explanation) that elucidates the concept of explanation in such a way that the instrumentalist may consistently agree that science contains explanations in the explicated sense, will not be an explication of any of the explicanda described here. (1973, 61–62)

Coffa's main object, in raising the instrumentalism-realism issue was to apply it to the received view. According to Hempel's account, a deductive-nomological explanation is an argument that contains essentially a lawful generalization. This law may be a generalization concerning observables only—for example, our old friend the ideal gas law. The subsumption of a fact under that generalization constitutes a satisfactory explanation on Hempel's account. If we go on to explain the ideal gas law in terms of the kinetic-molecular theory, that is a distinct explanation, and its existence does nothing to impair the status of the foregoing explanation. As the received view has it, there are legitimate scientific explanations that do not appeal to theories about unobservables; indeed, it is possible in principle (in a world very different from ours), though contrary to fact, for all legitimate explanations to be given wholly in terms of laws governing observable phenomena.¹⁴ Thus, it is possible in principle for the instrumentalist to embrace the received view without qualms. To Coffa's mind, this constitutes a severe shortcoming of any account of scientific explanation. I agree.

To see how these considerations apply to the ontic conception, let us focus upon causal explanation. The Humean tradition suggests that the world is full of constant conjunctions among observable phenomena, but that there is serious doubt as to whether there is anything in the external world that distinguishes ac-

cidental conjunctions from genuine causal relations. Hume seems to say that the distinction is in the mind—in the imagination. I do not believe that causality resides in the mind; moreover, I think it is a distinctly nontrivial matter to distinguish genuine cause-effect relations from accidental correlations or from cases in which the correlation is produced by a common cause. Furthermore, in many cases in which we seek a cause we do not even have a constant conjunction *among observables* to appeal to. For example, the nearly simultaneous illness of a number of people who had attended the American Legion convention in Philadelphia in 1976 could not be satisfactorily explained until the *Legionella* bacillus was identified, and its source in Philadelphia located.

As I said above, in "Why Ask, 'Why?'" I attempted to spell out in some detail the nature of the causal mechanisms that seem to exist in our world. They include causal processes, causal interactions (interactive forks), and conjunctive forks. Our casual observation of phenomena seldom reveals the causal mechanisms; indeed, careful scrutiny of the observable phenomena does not generally reveal their presence and nature. So our efforts at finding causal relations and causal explanations often—if not always—take us beyond the realm of observable phenomena. Such knowledge is empirical knowledge,¹⁵ and it involves descriptive knowledge of the hidden mechanisms of the world, but it does go beyond descriptive knowledge of the observable phenomena. There is no *logical* necessity in the fact that causal mechanisms involve unobservables; that is just the way our world happens to work.

As we have already noted, Coffa is a staunch defender of the ontic conception of scientific explanation, and his theory of explanation reflects this attitude. An explanation of any occurrence is a set of objective facts and relationships. For Coffa, what explains an event is whatever produced it or brought it about. Explanations are nomic dispositions of universal ($p = 1$) or less than universal ($p < 1$) strength. The linguistic entities that are often called 'explanations' are statements reporting on the actual explanation. Explanations, in his view, are fully objective and, where explanations of nonhuman facts are concerned, they exist whether or not anyone ever discovers or describes them. Explanations are not epistemically relativized, nor (outside of the realm of human psychology) do they have psychological components, nor do they have pragmatic dimensions. Since the mechanisms that operate in our world are frequently hidden, the true explanation (on Coffa's view) is often something whose existence the instrumentalist denies.

Traditionally, instrumentalism has been opposed to one sort of realism or another. The instrumentalist in physics recognizes that the molecular-kinetic theory of gases provides a useful tool for establishing regular relationships among such observables as temperature, pressure, volume, and mass, but denies that such things as molecules actually exist. The instrumentalist in psychology sees the function of psychological theories as providing relationships between observable

stimuli and observable responses, without appealing to such unobservable entities as feelings of hunger, anxiety, or pain.¹⁶ For the instrumentalist, our descriptive knowledge of the world is confined to knowledge of whatever is more or less directly observable. The instrumentalist cannot appeal to unobservables for purposes of explaining observed fact, for he or she denies that any such things exist.

The realist, in contrast, makes at least a rough demarcation between descriptive knowledge of observables and descriptive knowledge of unobservables. According to the realist, we can, in principle, have knowledge of observables by direct observation—though, in fact, much of our knowledge even of observables is rather indirect. The color, shape, size, and surface texture of a satellite of Pluto, for instance, can at present be known only by means of complex theoretical inferences. In addition, the realist claims, we can have descriptive knowledge of unobservables on the basis of theoretical inferences. Although some realists might maintain that our alleged knowledge of unobservables transcends the empirical realm, I claim that we have fully empirical knowledge of them. My thesis is that realism and empiricism are entirely compatible, for we can confirm or disconfirm theories about unobservables on the basis of observational evidence.¹⁷ The realist constructs explanatory theories that are intended and believed to make reference to unobservable entities. The realist asserts that such things as atoms, molecules, ions, subatomic particles, and microorganisms actually exist, and that we can explain a vast range of physical, chemical, and biological phenomena in terms of their behavior. As we shall see, this claim is sharply denied by van Fraassen, who is *not* an instrumentalist, but who shares a number of important views with philosophers of that persuasion.

For the proponent of the ontic conception of scientific explanation, realism provides a straightforward answer to the question of the distinction between descriptive and explanatory knowledge. Taking “description” in the narrower sense which includes only description of appearances, the realist can say that explanatory knowledge is knowledge of the underlying mechanisms—causal or otherwise—that produce the phenomena we want to explain. To explain is to expose the internal workings, to lay bare the hidden mechanisms, to open the black boxes nature presents to us.

The foregoing discussion of the relationship between descriptive and explanatory knowledge has focused almost exclusively on the intellectual value of scientific explanations. We should not forget that explanations have practical value as well. Finding scientific explanations of various types of occurrences often points to useful ways of controlling important features of our world. It may help to eliminate such undesirable events as epidemics and airplane crashes. It may help to bring about such desirable results as a greater healthy life span for humans. It may also help to alleviate superstitious fears.¹⁸ But, to return to our main theme, our discussion has left us with three apparently viable answers to the question of the intellectual value of scientific explanations. Such explanations enhance our un-

derstanding of the world. Our understanding is increased (1) when we obtain knowledge of the hidden mechanisms, causal or other, that produce the phenomena we seek to explain, (2) when our knowledge of the world is so organized that we can comprehend what we know under a smaller number of assumptions than previously, and (3) when we supply missing bits of descriptive knowledge that answer why-questions and remove us from particular sorts of intellectual predicaments. Which of these is *the* function of scientific explanation? None *uniquely* qualifies, I should say; all three are admissible. But not everyone would agree. And even among those who do agree, some will say that one of the three is fundamental, and the others have a distinctly derivative status.

4.4 The Pragmatics of Explanation

The most articulate and prominent anti-realist of the fourth decade is Bas van Fraassen. The full statement of his position appears in *The Scientific Image* (1980), but it had been anticipated to some extent in articles that appeared near the close of the third decade (1976, 1977). It should be carefully noted that van Fraassen is *not* an instrumentalist; his position is *constructive empiricism*. Like a sophisticated instrumentalist, van Fraassen recognizes that theories, which at least appear to make reference to unobservable entities, have played an indispensable role in the development of modern science, and he recognizes that in all likelihood they will continue to do so. The most fundamental difference between the instrumentalist and the constructive empiricist is that the former *denies* the existence of unobservables while the latter remains *agnostic* with respect to their existence.

According to van Fraassen, accepted scientific theories are *accepted as empirically adequate*, but they need not be *believed to be true*. Thus, he claims, when we accept a theory that seems to make reference to unobservables for various purposes—including use in giving explanations—we are committed to claiming that it yields true statements about observables, but we are not committed to claiming that what it says about unobservables is true. Likewise, however, we are not committed to claiming that what it says about unobservables is false. Van Fraassen's thesis is that one cannot be convicted of irrationality for disbelief in such things as molecules and one cannot be convicted of irrationality for belief in them. Regarding the name he has chosen for his position, he says, "I use the adjective 'constructive' to indicate my view that scientific activity is one of construction rather than discovery: construction of models that must be adequate to the phenomena, and not discovery of truth concerning the unobservable" (1980, 5).

In spite of his anti-realism, van Fraassen offers a theory of explanation that fits with his overall conception of the nature of science. Contrasting his view with the more traditional approaches (including, of course, the received view), he says,

The discussion of explanation went wrong at the very beginning when explanation was conceived of as a relationship like description: a relation between theory and fact. Really it is a three-term relation, between theory, fact, and context. No wonder that no single relation between theory and fact ever managed to fit more than a few examples! (1980, 156)

This statement contains a complete rejection of the conception upon which Coffa focused his attention, an ontic conception that located explanations in the external world and which is totally unavailable to instrumentalists.

The theory of explanation van Fraassen offers is *not* intended to be a theory only of scientific explanation; it should encompass other kinds of explanations as well. On his view, an explanation is *an answer to a why-question*. A scientific explanation is one that relies essentially on scientific knowledge. We might ask, for example, why Hitler invaded Russia during World War II. It is a historical question, and it calls for a historical answer. Whether or not history is a science, an answer to this question will be an explanation, and it will be within the purview of van Fraassen's theory.¹⁹ Similarly, one might ask why the cat is sitting in front of the door. The common-sense answer is that he wants to go outside. Again, van Fraassen's theory is meant to handle such questions and answers. History and common sense are, after all, closely related to science whether or not they qualify as parts of science. It is not too unreasonable to expect explanations in those domains to resemble scientific explanations.

There are, however, why-questions that do not seem to be calls for explanations. One might ask, in a time of grief, why a loved one had died. Such a question is not intended to evoke an explanation; it is a cry for sympathy. Exclamations and tears may constitute a far better response than would any factual proposition. Indeed, the grieving individual may be fully aware of the scientific explanation of the demise. Other why-questions seem best interpreted as requests for moral justification. The question has been raised in courts of law as to why a member of a minority group was admitted to medical school to the exclusion of some non-minority candidate whose qualifications were somewhat better. The point at issue is the ethical basis for that decision. It appears that many why-questions are not requests for explanations; consequently, to sustain the claim that explanations are answers to why-questions we would have to distinguish *explanation-seeking* why-questions from other kinds of why-questions. Van Fraassen does not undertake that task, but one might reasonably claim that contextual cues, which play a central role in his theory, should enable us to sort them out.

Given that not all why-questions seek to elicit explanations, the next question is whether all explanations—at least, all scientific explanations—are sought by means of why-questions. Clearly the answer to this question is negative. It has been suggested, however, that any request for a scientific explanation—no matter how it is actually formulated—can be appropriately rephrased in the form of a

why-question. Inasmuch as van Fraassen claims that explanations are answers to why-questions, he is patently committed to this view. A number of philosophers have denied it; for instance, Bromberger, William Dray, and Frederick Suppe have argued that some explanations are answers to questions of *how-possibly*. Dray had raised this issue (1957) and it is discussed by Hempel (1965, 428–30).

Consider a concrete example. There is an old saying that when a cat falls (from a sufficient height) it always lands on its feet. We know, however, that angular momentum is conserved. How is it then possible for a cat, dropped from an adequate height with zero angular momentum and with its legs pointing upward, to land on its feet? Your first reaction might be to suppose that the old saying is simply not true—that it is an old wives' tale. But that is not correct. Experiment has shown that the cat can twist its body in various ways while the net angular momentum remains zero to achieve the desired position upon landing.²⁰ A diver who does a twist, as distinguished from a somersault, achieves a similar feat. Hempel's response to such examples is to admit that there is a pragmatic difference between why- and how-possibly-questions, in that the person who poses a how-possibly question is under the mistaken impression that the occurrence is either physically impossible or highly improbable. The appropriate response is to expose the misapprehension and produce either a D-N or an I-S explanation of the phenomenon in question. The original question could thus have been rephrased, "Why did this event (which I initially regarded as impossible or highly improbable) occur?" (Hempel 1965, 428–30).

In (1984, 10) I also expressed the claim that all requests for scientific explanations can be formulated as why-questions, but I now suspect that it is mistaken. Hempel's response seems inadequate for two reasons. First, a how-possibly question does not require an actual explanation; any potential explanation not ruled out by known facts is a suitable answer. For example, a DC-9 jet airplane recently crashed upon takeoff at Denver's Stapleton Airport during a snowstorm. One peculiar feature of this accident is that the plane flipped over onto its back. There are many explanations of a crash under the circumstances, but I wondered how it could have flipped over. Two how-possibly explanations were mentioned in the news reports. One is that it encountered wing-tip turbulence from another airplane just after it became airborne. Another was suggested by the report of a survivor, who claimed that the plane was de-iced three times during its wait for departure, but that on the latter two of these occasions one wing, but not the other, was treated. If one wing had an accumulation of ice on its leading edge while the other did not, the difference in lift provided by the two wings might have been sufficient cause for the plane to flip over. As I write this paragraph I have not yet heard the final determination regarding the cause of this crash. Both potential explanations I have mentioned are satisfactory answers to the how-possibly question, but we do not know the correct answer to the why-question.

Second, improbable events do occur. Not long before the Denver crash, an-

other DC-9 crashed on take-off at Detroit's Metropolitan Airport. Investigators have concluded, I believe, that the pilot failed to extend the wing flaps for take-off. It is extremely unlikely that an experienced pilot would make such an error, that the co-pilot would fail to notice, and that the warning signal would fail to be sounded or would be ignored. But apparently that is what happened. Hempel suggests that we are obliged to find additional factors that would make the errors highly probable. It seems to me that, even if one insists on high probabilities for explanations that answer why-questions, no such thing is required for answers to how-possibly questions. It is sufficient to show that the probability is different from zero.

Still other explanations may be answers to *how-actually* questions. How did there come to be mammals (other than bats) in New Zealand? They were humans, who came in boats, and who later imported other mammals. This is, I think, a genuine scientific explanation. It is not an explanation of why they came; rather, it is an explanation of how they got there.

Having posted some caveats about simply identifying explanations as answers to why-questions, I shall now turn to a discussion of some of the details of van Fraassen's theory of explanation. Before doing so, it is worth noting that, although this issue is crucial for an advocate of the erotetic conception, it has little—if any—genuine significance for the proponent of the ontic conception. According to this latter conception the search for explanations is a search for underlying mechanisms, and the form of the question requesting them is not very important.

The development of van Fraassen's theory of why-questions and their answers in the fourth decade was greatly facilitated by progress in formal pragmatics and the publication, near the end of the third decade, of Belnap and Steel's *The Logic of Questions and Answers* (1976). Although this is a landmark work in erotetic logic, it contains hardly any treatment of why-questions. To begin, it is essential to realize that, as van Fraassen sets forth his theory, questions and answers are abstract entities. An answer is a proposition, and a given proposition can be expressed by means of many different declarative sentences. Moreover, a particular sentence, uttered on different occasions, can express different propositions. For any given sentence, the context determines which proposition it expresses. "I am here," for example, always expresses a true proposition, but the proposition it expresses depends on who utters the sentence, and where and when. Similarly, a given question may be expressed by many different interrogative sentences, and a particular interrogative may pose different questions on different occasions. When an interrogative sentence is uttered the context determines which question is being asked.

According to van Fraassen, we can think of the standard form of a why-question as

Why (is it the case that) P_k ? (Q)

where P_k states the fact to be explained (the explanandum phenomenon). Such a question can be identified with an ordered triple

$$Q = \langle P_k, X, R \rangle$$

where P_k is the *topic* of the question, $X = \{P_1, P_2, \dots, P_k, \dots\}$ is the *contrast class*, and R is the *relevance relation*. To take a familiar example, consider the question,

Why did the Bunsen flame turn yellow?

The topic is

The Bunsen flame turned yellow.

The contrast class is

The Bunsen flame remained blue (P_1)

The Bunsen flame turned green (P_2)

The Bunsen flame turned orange (P_3)

.

.

.

The Bunsen flame turned yellow (P_k)

.

.

.

The relevance relation R is the relation of cause to effect. Hempel's answer, we recall, is that a piece of rock salt was placed in the flame, rock salt is a sodium compound, and all sodium compounds turn Bunsen flames yellow.

A crucial feature of van Fraassen's account is its emphasis upon the fact that the same interrogative sentence—the same group of words—can express different questions. This can easily happen if different contrast classes are involved. He invites consideration of the interrogative

Why did Adam eat the apple?

By the inflection or emphasis of the speaker, or by other contextual clues, we might find that any of three different questions is being expressed. It might mean,

Why did *Adam* eat the apple?

where the contrast class = {Eve ate the apple, the serpent ate the apple, the goat ate the apple, etc.}. At the same time, it might mean

Why did Adam *eat* the apple?

where the contrast class = {Adam ate the apple, Adam threw the apple away, Adam gave the apple back to Eve, Adam fed the apple to the goat, etc.}. Also, it might mean

Why did Adam eat the *apple*?

where the contrast class = {Adam ate the apple, Adam ate the pear, Adam ate the pomegranate, etc.}. The context determines which is the appropriate contrast class.

Another feature of questions is that they generally come with *presuppositions*, and why-questions are no exception. The presupposition of Q is

- (a) P_k is true,
- (b) each P_j in X is false if $j \neq k$,
- (c) there is at least one true proposition A that bears the relation R to the ordered pair $\langle P_k, X \rangle$

where (a) and (b) taken together constitute the *central presupposition*. The canonical form of a *direct answer* to Q is

- (*) P_k in contrast to the rest of X because A.

The proposition A is known as the *core of an answer* to Q, because the direct answer would normally be abbreviated, "Because A." The following conditions must be met if (*) is to qualify as a direct answer to Q:

- (i) A is true.
- (ii) P_k is true.
- (iii) No member of X other than P_k is true.
- (iv) A bears relation R to $\langle P_k, X \rangle$.

The context in which a question is posed involves a body of background knowledge K. According to van Fraassen, two of the biggest problems faced by other theories of scientific explanation are *rejections of requests for explanation* and *the asymmetries of explanation*. Rejections are handled in van Fraassen's theory in terms of the presupposition. Unless K entails the truth of the central presupposition—namely, that the topic is true and that every member of the contrast class other than the topic is false—the question does not *arise* in that context. For example, the question "Why was Jimmy Hoffa murdered?" does not arise in my current knowledge situation, for to the best of my knowledge it has not been established that he was murdered—only that he disappeared. In addition, if K entails the falsity of (c)—that is, if K entails that there is no answer—the question does not arise. If, for instance, one asks why a given unstable nucleus decayed at some particular moment, we might answer that there is no reason why. Our best theories tell us that such things just happen. In other words, it is appropriate

to raise the question Q if we know that P_k is the one and only member of the contrast class that is true, and we do not know that Q has no answer.

If the question Q does not arise in a given context, we should reject it rather than trying to provide a direct answer. This can be done, as we saw in the preceding paragraph, by providing a corrective answer to the effect that some part of the central presupposition is not entailed by the body of knowledge K , or that K entails that there is no direct answer. If the question does arise, but we find out that (c) is false, then a corrective answer to that effect is appropriate. If the presupposition is completely satisfied, the request for a direct answer is legitimate, and if one is found, it constitutes an explanation of P_k .

There is, I believe, a profound difficulty with van Fraassen's theory centering on the relevance relation R . It can be put very simply, namely, that he imposes no restriction whatever on the nature of the relevance relation. He says explicitly that A is relevant to P_k if A bears relation R to P_k (1980, 143). But if R is not a bona fide relevance relation, then A is 'relevant' to P_k only in a Pickwickian sense. The difficulty can be posed in extremely simple terms. Formally speaking, a relation consists of a set of ordered pairs. Suppose we want to explain any fact P_k . Pick any arbitrary true proposition A . Let the relation R be the unit set of ordered pairs $\{ \langle A, \langle P_k, X \rangle \rangle \}$ – the set that has $\langle A, \langle P_k, X \rangle \rangle$ as its only member. "Because A " is an explanation of P_k .

Although I had studied van Fraassen's theory with some care between 1980 (when it was published) and 1985 (when the Minnesota NEH Institute was held), I had not noticed this difficulty. It emerged as a result of a discussion with Philip Kitcher, and we published a joint paper in which it was exhibited (1987). The problem was masked by a number of van Fraassen's informal remarks. For instance, at the outset of his exposition of his theory of why-questions, he says, "This evaluation [of answers] proceeds with reference to the part of science accepted as 'background theory' in that context" (1980, 141). Earlier, he had remarked that, "To ask that . . . explanations be scientific is only to ask that they rely on scientific theories and experimentation, not on old wives' tales" (1980, 129) and "To sum up: no factor is explanatorily relevant unless it is scientifically relevant; and among the scientifically relevant factors, context determines explanatorily relevant ones" (1980, 126). In conclusion, he says, "To call an explanation scientific, is to say nothing about its form or the sort of information adduced, but only that the explanation draws on science to get this information (at least to some extent) and, more importantly, that the criteria of evaluation of how good an explanation it is, are being applied using a scientific theory" (1980, 155–56). But in the formal account no restriction is imposed on the relation R .

To see the consequences of this lacuna, consider a concrete example. Suppose someone asks why John F. Kennedy died on 22 November 1963; this is the question $Q = \langle P_k, X, R \rangle$, where

P_k = JFK died 11/22/63	(topic)
X = {JFK died 1/1/63, JFK died 1/2/63, . . . JFK died 11/22/63, . . . JFK died 12/31/63, JFK survived 1963}	(contrast class)
R = astral influence	(relevance relation)

Suppose that the direct answer is

P_k in contrast to the rest of X because A ,

where A (the core of the answer) consists of a *true* description of the configuration of the planets, sun, moon, and stars at the time of Kennedy's birth. Suppose further that the person who supplies this answer has an astrological theory from which it follows that, given A , it was certain, or highly probable, that Kennedy would die on that day. We now have a why-question and an answer; the answer is an explanation. We must ask how good it is.

Van Fraassen's theory does not stop at the definition of an answer to a why-question; obviously, it must offer grounds for evaluating answers. We need to be able to grade explanations as better or worse. And he does provide criteria for this purpose; there are three. First, we must ask how probable the answer is in light of our background knowledge K . Second, we must ask to what extent the answer *favors* the topic vis-à-vis the other members of the contrast class. Third, we must ask how this answer compares with other available answers: (i) are any other answers more probable? (ii) do any other answers more strongly favor the topic? or (iii) do any other answers render this one wholly or partially irrelevant? On each of these criteria the astrological explanation gets highest marks. In the first place, since we have excellent astronomical records we have practical certainty regarding the celestial configuration at the time of Kennedy's birth; indeed, we have stipulated that A is true. In the second place, we must suppose that the astrologer can derive from A , by means of astrological theory, that Kennedy was sure to die on that day—or, at least, that the probability for his death on that day was much greater than for any other day of the year, and also much greater than the probability that he would live to see the beginning of 1964. Since this explanation, like any explanation, is given *ex post facto*, we must credit the astrologer with sufficient ingenuity to produce such a derivation. In the third place, no other

answer is better than A: (i) A is true; hence, no other answer could be more probable. Moreover, since, astrologically speaking, the heavenly configuration at the time of one's birth is the primary determinant of one's fate (ii) no other answer could favor the topic more strongly and (iii) no other facts could supersede answer A or render it irrelevant.

Consideration of the astrology example makes vivid, I hope, the fundamental problem van Fraassen's theory encounters with respect to the relevance relation. In our critique of that theory, Kitcher and I show formally that *any* true proposition A can be an indispensable part of an explanation of *any* topic P_k (with respect to a contrast class that contains P_k and *any* assortment of false propositions), and, indeed, that it gets highest marks as an explanation of P_k (1987, 319–22). Thus, it is natural to suggest that van Fraassen add one item to the list of presuppositions of Q, namely,

(d) R is a relevance relation.

But when we attempt to impose such a condition, we find ourselves in a mare's nest of difficulties. The problem is to characterize in a general way what constitutes a suitable relevance relation, recognizing, of course, that there may be more than one.

Consider, for example, the relation of logical deducibility. That this relation is not satisfactory has been known since antiquity; Aristotle pointed out that some demonstrations provide understanding while others do not. Hempel recognized from the outset that demonstrations like

Horace is a member of the Greenbury School Board.

All members of the Greenbury School Board are bald.

Horace is bald.

cannot qualify as a bona fide explanation of Horace's baldness. Since the beginning of the first decade in 1948 he has insisted that the demonstration must contain essentially at least one law-statement among its premises. But we have seen what a vexed question it is to distinguish lawlike from nonlawlike statements.

We have also seen that, even if the problem of lawlikeness can be handled, a problem about asymmetries arises. Recalling one of our standard counterexamples, we have two valid deductive arguments with suitable lawful premises, one of which seems clearly to provide an explanation, the other of which seems to most of us not to. From the height of the flagpole in conjunction with the elevation of the sun and laws of propagation of light, we can deduce the length of the shadow it casts. This sounds like a good explanation. From the length of the shadow and other premises of the aforementioned sort (excluding, of course, the height of the flagpole), we can deduce the height of the flagpole. This does not sound like a good explanation. Hence, the problem of asymmetries.

Van Fraassen has maintained, as we have already noted, that the two chief problems inherited from the traditional accounts of explanation that he wants to solve are the problem of rejections and the problem of asymmetries. We have already discussed his solution to the problem of rejection of the explanatory questions, and found no fault with it. His solution to the problem of asymmetries is another story—indeed, it is the fable of “The Tower and the Shadow” (1980, 132–34).

Van Fraassen’s treatment of the asymmetries is to show that certain why-questions usually arise in certain typical contexts in which a standard sort of answer is satisfactory. According to van Fraassen’s story, when first he asks why the shadow of the tower is so long, he is told that it is cast by a tower of a certain height; in addition, his host, the Chevalier, adds that the tower was built to that height on that particular spot for certain historical reasons. That is *his* explanation, but later in the tale we learn that it is false. The servants have a different, and more accurate, explanation of the position and height of the tower. Carefully taking the contextual factors into account, we discover that the correct answer to the question, “Why is the tower that tall and located in that place,” is that it casts a shadow long enough to cover a certain spot on the terrace, where the Chevalier had murdered a servant girl in a fit of jealous rage, at a certain time of day. Here we have an admissible why-question and a suitable direct answer. We are tempted to remonstrate, as I did (1984, 95), that it was the antecedent desire of the Chevalier to have a shadow that long that explains the height of the tower. But van Fraassen seems to be maintaining that the answer given is legitimate. The topic P_k of the question is that the tower stands at a particular place and is 175 feet tall. The contrast class X consists of a series of statements about towers of different heights located at various locations in the vicinity. And the relation R is a relation of intentional relevance. Our remonstrance was based on the belief that this relation is not a suitable explanatory relevance relation. As an advocate of a causal/mechanical conception of scientific explanation, I am not prepared to admit that effects explain their causes, even where conscious purposeful behavior is involved.

It is clear from the whole tenor of van Fraassen’s account that he is not proposing an ‘anything goes’ sort of theory. He is *not* suggesting that any answer to a why-question that happens to satisfy a questioner is a satisfactory explanation. Consequently, it seemed to Kitcher and me, van Fraassen’s pragmatic theory cannot escape precisely the kinds of problems concerning objective explanatory relevance relations with which other more traditional theories—such as the received view or the statistical-relevance approach—had to struggle. Hence, when I proposed above the addition of presupposition

(d) R is a relevance relation

I was opening up the whole question of what constitutes a satisfactory explanatory relevance relation.

Consider, for example, the relation of statistical relevance, which I once regarded as the key explanatory relation. We suppose that the person who raises a given why-question has a prior probability distribution over the members of the contrast class. When an appropriate answer is given, it results in a different probability distribution over the same set of alternatives. In the S-R model I required only that the probabilities change; van Fraassen's theory requires that the topic be favored—that is, that it be elevated relative to its rivals in the contrast class. What kinds of probabilities are these? Coffa and I took propensities and frequencies, respectively, as the appropriate interpretations. In a pragmatic theory of explanation, because of its emphasis upon the knowledge situation and the context, it is natural to think of epistemic or personal probabilities. Given the well-known difficulties with the former, van Fraassen seems to prefer the latter.

Whatever sort of probability is involved, a basic problem—one we have confronted previously—arises. When an explanation is sought we already know that the explanandum-phenomenon has occurred. For van Fraassen, the truth of the topic is the first presupposition of a why-question. Moreover, the second presupposition of the why-question is that all of the other members of the contrast class are false. The why-question does not arise unless these presuppositions are fulfilled. But if this presupposition is fulfilled, the prior probability distribution with respect to the body of knowledge K is one for the topic and zero for all other members of the contrast class. No change in this probability distribution could possibly favor the topic. So van Fraassen proposes that we must cut back our body of knowledge K to some proper part $K(Q)$ that is appropriate for question Q (1980, 147). The problem of deciding what information $K(Q)$ should contain is precisely the problem Hempel faced in connection with the requirement of maximal specificity. It is also the problem with which I had to deal in characterizing objective homogeneity. It turns out, then, that this cluster of traditional problems is not evaded by van Fraassen's pragmatic account.

The appeal to personal probabilities in this context gives rise to another serious difficulty. Return to the example of John F. Kennedy's assassination. To the sincere believer in astrology the configuration of heavenly bodies at the time of Kennedy's birth is highly relevant to his death on that particular fateful day in 1963. Acquiring that information will produce a redistribution of personal probabilities strongly favoring the topic. Believing, as we do, that there is no objective relevance between the celestial configuration at the time of Kennedy's birth and the occurrence of his death on a particular day, we need to block such explanations. Unless we can impose the demand for objective relevance relations, we cannot arrive at a satisfactory characterization of scientific explanation. As many philosophers have insisted, we need to appeal to objective nomic relations, causal relations, or other sorts of physical mechanisms if we are to provide adequate scientific explanations.

Philosophers have long recognized that scientific explanation has pragmatic

dimensions. It is obvious that various features of actual explanations depend upon context. Hempel was aware from the beginning that the individual who answers an explanation-seeking why-question standardly omits parts of the explanation that are already well known to the questioner. I have emphasized the fact that one frequently has to clarify the question in order to ascertain what explanation is being sought. As we saw in the second decade, the ordinary language philosophers placed great emphasis upon pragmatic considerations. The theory of the pragmatics of explanation given by van Fraassen in *The Scientific Image* is highly illuminating and is, I believe, the best that has been given to date. It must be emphasized, however, that he has not succeeded in showing that all the traditional problems of explanation can be solved by appealing to pragmatics. In that sense he has *not* provided a pragmatic theory of explanation. The problems concerning the nature of laws, and those concerning the nature of causality, have not been circumvented by pragmatic considerations.

Another important representative of the erotetic version of the epistemic conception is Peter Achinstein, whose view is articulated in great detail in *The Nature of Explanation* (1983). His theory differs significantly from those of both Bromberger and van Fraassen, but it is much closer in spirit to that of Bromberger. This can be seen most clearly in the emphasis that both Achinstein and Bromberger place on the linguistic analysis of English usage—something van Fraassen does hardly at all. Whereas van Fraassen simply announces that explanations are answers to why-questions, Achinstein invests considerable time and effort in clarifying the usage of “explanation” and closely related terms. Along with Bromberger, he denies that all explanations are answers to why-questions. Nevertheless, as we shall see, questions—not just why-questions—and their answers play a fundamental role in this theory.

Achinstein points out that “explanation” may refer either to a *process* or a *product*. The process is a linguistic performance; someone explains something to someone by uttering or writing statements. The product is the content of the linguistic performance. The linguistic performance itself involves an intention on the part of person producing the explanation: “Explaining is what Austin calls an illocutionary act. Like warning and promising, it is typically performed by uttering words in certain contexts with appropriate intentions” (1983, 16). According to Achinstein, this process concept of explanation is primary. When we try to characterize the product, we must take account of the intention (or illocutionary force) of the explanation, for the same set of words can be used either to explain or do other sorts of things. A physician might explain John’s malaise this morning by saying, “He drank too much last night.” John’s wife might use the same words to criticize his behavior. His wife’s speech act is not an explanation; consequently, what she produced is not an explanation. This same pragmatic consideration arose, incidentally, in the 1948 Hempel-Oppenheim essay for, given the explanation/prediction symmetry thesis, an argument may function in one context

as an explanation and in another as a prediction. To deal explicitly with this aspect of explanation, Achinstein adopts what he calls an *ordered pair view* of explanation: an explanation in the product sense is an ordered pair $\langle x, y \rangle$ in which x is a specified type of proposition and y is a type of speech act, namely, explaining. On this view, y retains the intention involved in the process of explanation (1983, 85–94). Achinstein refers to his account as *the illocutionary theory*.

It should be noted that the foregoing considerations are designed to clarify the notion of explanation without qualification. Up to this point, Achinstein makes no attempt to characterize correct explanations, good explanations, or scientific explanations. It is not until page 117, almost one-third of the way through a fairly large book, that scientific explanation comes up for serious consideration. The philosopher of science who is impatient to get to that topic cannot begin there, however, for the preliminaries are used extensively in the subsequent discussions.

In his preliminary formulations, Achinstein presents two aspects of explanation:

- (1) If S explains q by uttering u , then S utters u with the intention that his utterance of u render q understandable (1983, 16);
- (2) If S explains q by uttering u , then S believes that u expresses a proposition that is a correct answer to Q (1983, 17).

If Q is a why-question (which is, I repeat, for Achinstein only one among many sorts of explanation-seeking questions), then q is what van Fraassen called the *topic* of the question. To use one of Achinstein's examples, if Q is "Why did Nero fiddle?" then q is "Nero fiddled." According to (1), whatever answer is given is intended to make q understandable, but since explanation and understanding are such closely related concepts, it is necessary for Achinstein to say something about what constitutes understanding. He offers the following necessary condition:

A understands q only if there exists a proposition p such that A knows of p that it is a correct answer to Q , and p is a complete content-giving proposition with respect to Q . (Here p is a proposition expressed by a sentence u uttered by A.) (1983, 42)

He later suggests that it is a sufficient condition as well (1983, 57). Space does not permit a full statement of what constitutes a complete content-giving proposition with respect to a question; the details are spelled out in Achinstein's book. But the crucial point can be raised by considering a special case—one of his examples. A straightforward complete content-giving proposition with respect to the question, "Why did Nero fiddle?" is "The reason Nero fiddled is that he was happy." Given that this is a complete content-giving proposition with respect to that question, A understands Nero's fiddling iff A knows that "The reason Nero

fiddled is that he was happy" is a *correct answer* to the question, "Why did Nero fiddle?"

This view of explanation seems seriously question-begging. We may raise essentially the same question with regard to Achinstein's theory as we did concerning van Fraassen's: what objective relationship must obtain between the fact that Nero was happy and the fact that he fiddled to make "The reason Nero fiddled is that he was happy" a correct answer? How must that relationship differ from the relationship between the fact that Caesar was assassinated on the Ides of March and the fact that Nero fiddled? These questions have fundamental importance; to see this it will be useful to make a direct comparison with Hempel's theory of deductive-nomological explanation.

According to Hempel's theory, if "Nero was happy" is part of the explanans of "Nero fiddled," then "Nero was happy" must be true. Hempel's theory of explanation specified the relationship that must exist between the facts if one is to be (part of) an explanation of the other. According to the D-N model, the statement "Nero was happy" must be a premise of a valid deductive argument having "Nero fiddled" as its conclusion, and including essentially at least one other premise stating a lawful regularity. This argument must fulfill the Hempel-Oppenheim empirical condition of adequacy, namely, that all of its premises be true. Given the fulfillment of these conditions, we are then authorized to accept the claim that (at least part of) the reason Nero fiddled is that he was happy. It was *not* part of the empirical condition of adequacy to determine that "The reason Nero fiddled is that he was happy" is true. The whole idea of the Hempel-Oppenheim theory was to provide conditions under which it is correct to make such claims as "The reason Nero fiddled is that he was happy." They do not require us to assess the truth of such a statement to ascertain whether or not we have a correct explanation. Achinstein is clearly aware that this line of argument may be brought against his theory, and he attempts to rebut it (1983, 71–72), but I am not convinced that his defense is successful. I think he ends up—like Bromberger and van Fraassen—lacking an adequate characterization of the kinds of objective relevance relations required in sound scientific explanations.

Despite my skepticism regarding the illocutionary theory of explanation, there are, it seems to me, several especially illuminating features of Achinstein's treatment of scientific explanation. First, he distinguishes carefully between *correct* explanations and *good* explanations. An explanation $\langle x, y \rangle$ is a correct explanation if the first member of that ordered pair is a true statement. Of course, the fact that the ordered pair is an explanation imposes other qualifications on x . However, for any number of pragmatic reasons, a correct explanation may not be a good explanation. It may be unsuitable to the knowledge and abilities of the listeners, or lacking in salience with respect to their interests. To deal with evaluations of explanations over and above correctness, Achinstein introduces the idea of a set of instructions to be followed in constructing explanations (1983, 53–56).

Such instructions could be of a wide variety of kinds. One might be the Hempelian instruction that the explanation must include essentially at least one law. Another might be an instruction to give a microphysical explanation. Still another might be the instruction to give a causal explanation.

Another important feature is Achinstein's contention that there is no single set of universal instructions that will suffice to judge the merits of scientific explanations at all times and in all contexts. Instructions that are suitable in one context may be quite unsuitable in another. He offers a characterization of appropriate instructions as follows:

I is a set of appropriate instructions for an explainer to follow in explaining *q* to an audience iff *either*

- a. The audience does not understand *q* in a way that satisfies I, and
- b. There is answer to *Q* (the question that elicits an explanation of *q*), that satisfies I, the citing of which will enable the audience to understand *q* in a way that satisfies I, and
- c. The audience is interested in understanding *q* in a way that satisfies I, and
- d. Understanding *q* in a way that satisfies I, if it could be achieved, would be valuable for the audience;

or

It is reasonable for the explainer to believe that a-d are satisfied. (113, slightly paraphrased)

Employing several important historical examples, Achinstein argues that there is no set of instructions that is universally appropriate for science (1983, 119–56).

The discussion of universal instructions leads naturally into another main feature of Achinstein's theory, namely, a consideration of the possibility of formal models of explanation such as the D-N, I-S, S-R. Again the conclusion is negative. Achinstein states two requirements which, he believes, motivate the "modelists." The first is the "No-Entailment-By-Singular-Sentence" (or NES) requirement. According to this requirement, no correct explanation of a particular occurrence can contain, in the explanans, any singular sentence or finite set of singular sentences that entail the explanandum (1983, 159). The second requirement he calls "the *a priori* requirement." According to this requirement, "the only *empirical* consideration in determining whether the explanans correctly explains the explanandum is the truth of the explanans; all other considerations are *a priori*" (1983, 162). His strategy in arguing that there cannot be models of explanation, in the traditional sense, is that any model that satisfies one of these requirements will violate the other (1983, 164–92). Even though I have spent a great deal of effort in elaborating the S-R model, this view of Achinstein's is one with which I happen to agree. I now think that an adequate scientific explanation identifies the mechanisms by which the explanandum came about. Consequently, what constitutes a suitable scientific explanation depends on the kinds of mechanisms—

causal or noncausal—that are operative in our world. This is an issue that cannot be settled a priori. Achinstein's arguments against the possibilities of models of explanation are far more precise and detailed than mine.

4.5 Empiricism and Realism

Van Fraassen's seminal book, *The Scientific Image*, has spawned a great deal of discussion, most of which has been directed toward his rejection of scientific realism. Churchland and Hooker, *Images of Science* (1985), is an important collection of critical essays, accompanied by van Fraassen's detailed replies. None of these essays focuses primarily on his treatment of scientific explanation, though many are tangentially relevant to it.²¹ Since, however, the realism issue has direct bearing on the nature of theoretical explanation, we must devote some attention to it. To carry out this discussion I will accept van Fraassen's claim—which I believe to be sound—that there is a viable distinction between the observable and the unobservable. Although the dividing line may not be sharp, there are clear cases of observables (e.g., sticks and stones) and clear cases of unobservables (e.g., atoms, electrons, and simple molecules).

The first sentence in the first chapter of *The Scientific Image* is striking: "The opposition between empiricism and realism is old, and can be introduced by illustrations from many episodes in the history of philosophy" (1980, 1). It formulates an assumption that goes unquestioned throughout the rest of the book—namely, that it is impossible to have empirical evidence that supports or undermines statements about objects, events, and properties that are not directly observable by the unaided normal human senses. This assumption should not, I think, go completely unchallenged. Indeed, I believe it is false. However that may be, it raises what I take to be the *key question* for scientific empiricism (W. Salmon 1985).

Let me illustrate the point by means of a simple example. I own a copy of the Compact Edition of the Oxford English Dictionary. It contains print of various sizes. I can read the largest print on the title page with my naked eye, but everything else is blurry without my eyeglasses. When I put on the spectacles I can easily read some of the larger print within the books. The use of corrective lenses does not take us beyond the realm of the directly observable; their effect is to restore normal vision, not to extend it beyond the normal range. The spectacles enable me to see the things I could have seen in my early teens without their aid. But even with their aid much of the print is blurry. With the aid of a magnifying glass (which comes with the set) I can read even the entries in smallest type. Also, with the aid of the magnifying glass, I can see marks of punctuation that were completely invisible to me without it. I claim that I have established the existence of an entity that is not directly observable; I have established a statement to the effect that there is, at a particular place on a given page, an ink spot too small

to be detected with the normal unaided human sense of sight. Moreover, I cannot feel it, smell it, hear it, or taste it.

It is important to note that, when I view through the magnifying glass print that I *can* read without it, the letters, words, and marks of punctuation that I see are the same. They simply appear clearer and larger. When I view smaller print with the magnifying glass, the forms I see—letters, words, marks of punctuation—make sense. I see bona fide words, and I read appropriate definitions of them. The words I read appear in their correct places. The same is true of the dot that I could not see at all without the magnifying glass. When it is made visible by the glass, it appears in a syntactically correct place. Moreover, although I confess that I have not performed the experiment, I have complete confidence that a comparison of the entries in the compact edition with those of the unreduced editions would reveal an identity between what is seen without the magnifying glass in the larger edition with what is seen with the aid of the magnifying glass in the compact edition.

Evidently, many different experiments of the type just described can be conducted with a variety of lenses, and on their basis we can establish a theory of geometrical optics. A number of fundamental facts of geometrical optics were known in antiquity and medieval times, and Snell's law was proposed in 1621. This theory is completely empirical even in the narrow sense van Fraassen adopts. With geometrical optics we can develop the theories of the telescope and the microscope. These theories are readily confirmable by means of experiment. Telescopes, for example, can be used to view from a distance terrestrial objects that can be approached for viewing at close range. It is interesting that van Fraassen regards as observable celestial objects—such as the moons of Jupiter—that can be seen from earth only with the aid of a telescope. We can, in principle, travel closer to them and see them with the naked eye. More interesting still is the fact that he takes objects remote from us in time—such as dinosaurs—to be observables. Time-travel into the past is something NASA has not achieved even in a small way. In contrast, he considers objects that can be seen only with the aid of a microscope as unobservables. There is no other place to go to get a better vantage point.

To substantiate my own claim that we can have empirical knowledge of unobservables, I attempted to spell out the sort of inference that underlies the transition from observation to conclusions about unobservables (1984, 231–34). Shortly thereafter I put the matter in these terms:

When David Hume stated that all of our reasonings about unobserved matters of fact are based upon the relation of cause and effect he was, I suspect, almost completely correct. One notable exception is induction by simple enumeration (or some other very primitive form of induction). As remarked above, I am assuming for purposes of this discussion that some sort of primi-

tive induction is available; I shall not attempt to characterize it or justify it in this context. The type of argument required in connection with microscopic observation is, I think, causal. It is a rather special variety of causal inference that is also analogical. This particular sort of causal/analogical argument is, in my view, quite powerful. To employ it, we must assume that we already have knowledge of cause-effect relations among observables—e.g., that hitting one's thumb with a hammer causes pain, that drinking water quenches thirst, and that flipping a switch turns off the light. Such relations can be established by Mill's methods and controlled experiments. Neither the instrumentalist's nor the constructive empiricist's account of science can get along without admitting knowledge of such relations. Using relations of this sort, we can now schematize what I suspect is the basic argument enabling us to bridge the gap between the observable and the unobservable. It goes something like this:

It is observed that:

An effect of type E_1 is produced by a cause of type C_1 .

An effect of type E_2 is produced by a cause of type C_2 .

.

.

.

An effect of type E_k occurred.

We conclude (inductively) that:

A cause of type C_k produced this effect of type E_k .

The particular application of this argument that interests us is the case in which $C_1, C_2, \dots, C_k$ are similar in most respects except size. Under these circumstances we conclude that they are similar in causal efficacy. (1985, 10–11)

The foregoing argument connects quite directly with an analysis of microscopic observation presented by Ian Hacking just after the publication of van Fraassen's *The Scientific Image*. To learn something about the scientific use of microscopes Hacking did something quite extraordinary for a philosopher. He actually went to a laboratory where the use of microscopes is essential to the research being conducted, and he learned how to use a variety of types. The research involved observation of dense bodies in red blood cells, and it employed a standard device known as a *microscopic grid*. "Slices of a red blood cell are fixed upon a microscopic grid. This is literally a grid: when seen through a microscope one sees a grid each of whose squares is labelled with a capital letter" (1981, 315). Making reference to the microscopic grid, he then addresses the issues raised by van Fraassen:

I now venture a philosopher's aside on the topic of scientific realism. Van Fraassen says we can see through a telescope because although we need the telescope to see the moons of Jupiter when we are positioned on earth, we

could go out there and look at the moons with the naked eye. Perhaps that fantasy is close to fulfillment, but it is still science fiction. The microscopist avoids fantasy. Instead of flying to Jupiter he shrinks the visible world. Consider the grid that we used for re-identifying dense bodies. The tiny grids are made of metal: they are barely visible to the naked eye. They are made by drawing a very large grid with pen and ink. Letters are neatly inscribed by a draftsman at the corner of each square on the grid. Then the grid is reduced photographically. Using what are now standard techniques, metal is deposited on the resulting micrograph. . . . The procedures for making such grids are entirely well understood, and as reliable as any other high quality mass production system.

In short, rather than disporting ourselves to Jupiter in an imaginary space ship, we are routinely shrinking a grid. Then we look at the tiny disk and see exactly the same shapes and letters as were drawn in the large by the first draftsman. It is impossible seriously to entertain the thought that the minute disk, which I am holding by a pair of tweezers, does not in fact have the structure of a labelled grid. I know that what I see through the microscope is veridical because we *made* the grid to be just that way. I know that the process of manufacture is reliable, because we can check the results with any kind of microscope, using any of a dozen unrelated physical processes to produce an image. Can we entertain the possibility that, all the same, this is some kind of gigantic coincidence[?]? Is it false that the disk is, in fine, in the shape of a labelled grid? Is it a gigantic conspiracy of 13 totally unrelated physical processes that the large scale grid was shrunk into some non-grid which when viewed using 12 different kinds of microscopes still looks like a grid? (1981, 316–17)

To avoid possible misinterpretation I should report that Hacking does not mean to use his argument to support the kind of wholesale realism I argued for in “Why Ask, ‘Why?’?” but he does offer a strong argument for the conclusion that we can, with the aid of microscopes, have knowledge of objects and properties not visible to the naked eye.²²

It seems to me that we can usefully distinguish between direct and indirect observation, where direct observation is accomplished by the use of unaided normal human senses, and indirect observation is accomplished by the use of instruments, such as the microscope and the telescope, that extend the range of the senses. I consider my simple argument, based in part on the development of geometrical optics, and Hacking’s argument, based on sophisticated microscopy, strong arguments to the effect that we can have indirect observational knowledge of objects and properties that are not directly observable. In addition, I want to claim, our knowledge of unobservables can be extended even further by appealing to appropriate theoretical considerations. That was the upshot of the discus-

sion of Perrin's argument concerning the ascertainment of Avogadro's number N and the issue of molecular reality in §4.2 above. Perrin studied the behavior of Brownian particles that could be observed using microscopes available in his day. He used his *indirect observations* of these entities as a basis for *inferring* the existence of simple molecules that are much too small to be viewed microscopically. It is mildly amusing that Perrin cited 13 independent methods of ascertaining N and that Hacking refers to 13 independent physical processes in dealing with the microscopic grid. The number 13 appears to be especially unlucky for antirealists.

One problem that has traditionally been associated with scientific realism is the problem of the meaning of theoretical terms. That cannot be the crucial problem, for we have been entertaining claims about unobservable entities without using any esoteric theoretical vocabulary. Since we do successfully describe the things we directly observe, our language must contain an observational vocabulary. However we might characterize it in general, "ink spot," "page," and "smaller than" are surely terms within it. Recall, in this connection, Perrin's work with tiny spheres of gamboge—*much* too small to be observed directly—which can be described entirely within our observational vocabulary.

I realize, of course, that there exists a broad range of philosophical opinion on the matter of scientific realism. I have recapitulated the argument that I find compelling. Other philosophers—e.g., Arthur Fine and Larry Laudan—join van Fraassen in rejecting realism, but for reasons quite different from van Fraassen's. Still others—e.g., Richard Boyd and Ernan McMullin—embrace realism, but appeal to different kinds of arguments to support their views.²³ We cannot escape the realism issue, for it is crucial to the debate between Coffa's highly realistic ontic conception of scientific explanation and van Fraassen's highly pragmatic erotetic approach.

4.6 Railton's Nomothetic/Mechanistic Account

In the first year of the fourth decade Peter Railton's first published article on scientific explanation appeared (1978). It is addressed chiefly to Hempel's I-S model, and it embodies an attempt to provide an account of probabilistic explanation that avoids what Railton regards as the two most troublesome aspects of Hempel's model—namely, epistemic relativization and the requirement of maximal specificity. This article can be viewed, in part, as a further development of some of the ideas expressed in Jeffrey (1969). With Jeffrey, Railton rejects the thesis of the received view that all explanations are arguments, but he goes further than Jeffrey in this regard. Whereas Jeffrey had admitted that in some "beautiful cases" (where the difference between the actual probability and unity is so small as to "make no odds") statistical explanations can be arguments, Railton argues that (practical considerations aside) there is no theoretical difference between the

beautiful cases and the unbeautiful. In none of the kinds of cases Hempel treated as inductive-statistical, he claims, should the explanation be construed as an argument. Railton also agrees with Jeffrey in maintaining that, for ineluctably statistical phenomena, the key to explanation lies in understanding the stochastic mechanism by which the occurrence came to pass, not in finding some way to render the event nomically to be expected. In this connection, both Jeffrey and Railton agree that, where some results of a given stochastic process are probable and others improbable, we understand the improbable just as well as we understand those that are highly probable. With Jeffrey, consequently, Railton rejects the high-probability requirement. Railton goes far beyond Jeffrey, however, in spelling out the details of a model of scientific explanation that embodies all of these features.

Railton chooses to elaborate what he calls a *deductive-nomological model of probabilistic explanation*—or *D-N-P model*—in terms of an example of an event that has an extremely small probability, namely, the alpha decay of a nucleus of uranium²³⁸. Since the mean-life of this radionuclide is 6.5×10^9 years, the probability that such an atom would decay within a specific short period of time is almost vanishingly small—but not quite vanishing, for such decays do occur. Our theory enables us to calculate the probability p that such a decay will occur within a short time interval Δt . Suppose u , a particular nucleus of this sort, has emitted an alpha-particle during such an interval. Then, we can set up the following deductive argument:

- (2) (a) All nuclei of U^{238} have probability p of emitting an alpha-particle during any interval of length Δt , unless subjected to environmental radiation.
 - (b) Nucleus u was a nucleus of U^{238} at time t and was subjected to no environmental radiation during the interval $[t, t + \Delta t]$.
-
- (c) Nucleus u had a probability p of emitting an alpha-particle during the interval $[t, t + \Delta t]$. (Railton 1978, 214, slightly paraphrased)

In this argument Railton, like Coffa, construes the probabilities as single-case propensities, and he regards premise (a) as a lawful generalization. He assumes, moreover, that (a) is an irreducibly probabilistic law, and that it incorporates all probabilistically relevant factors. He recognizes that (2) appears to be an explanation of the fact that u had a probability p of decaying within the specified time interval; he maintains, however, that it can be supplemented in a way that transforms it into an explanation of the fact that u actually decayed. The first addition is “a derivation of (2a) from our theoretical account of the mechanism at work in alpha-decay” (1978, 214)—from the theory of quantum-mechanical tunneling. This is, I take it, another deductive argument employing the Schrödinger wave

equation and such facts as the atomic weight and atomic number of ${}_{92}\text{U}^{238}$ as its premises. It would qualify as one of Hempel's D-S explanations.

Railton's approach is strongly mechanistic, as can be seen clearly by contrasting his attitude toward these two arguments with Hempel's attitude. For Hempel, we recall, to explain a particular fact it is sufficient to subsume it under a law. Such explanations are complete. If one wants an explanation of a law that entered into the first explanation, it can be supplied by deriving that law from more general laws or theories. The result is another explanation. The fact that a second explanation of this sort can be given does nothing to impugn the credentials of the first explanation. Railton's view is quite different. According to him—and in this I heartily agree—explanation involves revealing the mechanisms at work in the world. Mere subsumption of phenomena under generalizations does not constitute explanation. Explanation involves understanding *how the world works*. For Railton, then, the quantum mechanical explanation of the probabilistic decay law is an integral and indispensable part of the explanation of the decay that occurred.

To transform this pair of arguments, which still appear to explain the probability of the decay of *u*, into an explanation of the actual decay of *u*, we are asked to supply a "parenthetic addendum to the effect that *u* did alpha-decay during the interval" in question (1978, 214). These three components, taken together, though they do not constitute an argument or a sequence of arguments, do constitute an *explanatory account*:

- (3) A derivation of (2a) from our theoretical account of the mechanism of alpha decay.

The D-N inference (2).

The parenthetic addendum.

A D-N explanation and a D-S explanation (or two D-N explanations, if we continue to consider D-S a subtype of D-N) serve as the core of the explanatory account, but they do not comprise the whole explanation. The parenthetic addendum is also required. If the parenthetic addendum were taken as an additional premise for an argument, the explanation would be vitiated by becoming trivially circular.

If the explanatory account is not an argument, Railton realizes, many readers are going to wonder about its explanatory status:

Still, does (3) explain why the decay took place? It does not explain why the decay *had to* take place, nor does it explain why the decay *could be expected to* take place. And a good thing, too: there is no *had to* or *could be expected to* about decay to explain—it is not only a chance event, but a very improbable one. (3) does explain why the decay *improbably* took place, which is how it did. (3) accomplishes this by demonstrating that there existed at the time a

small but definite physical possibility of decay, and noting that, by chance, this possibility was realized. The derivation of (2a) that begins (3) shows, by assimilating alpha-decay to the chance process of potential barrier tunneling, how this possibility comes to exist. If alpha-decays are chance phenomena of the sort described, then once our theory has achieved all that (3) involves, it has explained them to the hilt, however unsettling this may be to *a priori* intuitions. To insist upon stricter subsumption of the explanandum is not merely to demand what (alas) cannot be, but what decidedly should not be: sufficient reason that one probability rather than another be realized, that is, chances without chance. (1978, 216)

Railton argues carefully that his characterization of probabilistic explanation does escape epistemic relativization, and that it has no need for a requirement of maximal specificity. The basic reason is that, by construing (2a) as a (putative) law, he claims that it is simply false if it fails to be maximally specific. If there are further factors relevant to the occurrence of the decay, which have not been included in the alleged law, it is false. Moreover, the kind of maximal specificity to which he refers is fully objective; it is just as objective as my objectively homogeneous reference classes. Inasmuch as Railton adopts a *single case propensity interpretation* of probability, he has no need for reference classes, but the maximal specificity of his laws is strictly analogous to objective homogeneity for a frequentist (see the discussion of propensities in §3.3 above).

In the introductory sections of his 1978 paper, Railton offers some extremely compact remarks about his approach to explanation in general, not just probabilistic explanations of particular facts. Regarding his mechanistic orientation, he says,

The goal of understanding the world is a theoretical goal, and if the world is a machine—a vast arrangement of nomic connections—then our theory ought to give us some insight into the structure and workings of the mechanism, above and beyond the capability of predicting and controlling its outcomes. . . . Knowing enough to subsume an event under the right kind of laws is not, therefore, tantamount to knowing the *how* and *why* of it. As the explanatory inadequacies of successful practical disciplines remind us: explanation must be more than potentially-predictive inferences or law-invoking recipes. (1978, 208)

Some of the mechanisms are, of course, indeterministic.

The D-N probabilistic explanations to be given below do not explain by giving a deductive argument terminating in the explanandum, for it will be a matter of chance, resisting all but *ex post facto* demonstration. Rather, these explanations subsume a fact in the sense of giving a D-N account of the chance mechanism responsible for it, and showing that our theory implies the exis-

tence of some physical possibility, however small, that this mechanism will produce the explanandum in the circumstances given. I hope the remarks just made about the importance of revealing mechanisms have eased the way for an account of probabilistic explanation that focuses on the indeterministic mechanisms at work, rather than the “*nomic* expectability” of the explanandum. (1978, 209)

The views expressed by Railton in this early article are ones I find highly congenial. There may, however, be a difference in our attitudes toward causal explanation. Like Jeffrey before him, Railton discusses the old chestnut of the falling barometric reading and the storm:

- (S) The glass is falling.
Whenever the glass falls the weather turns bad.
-

The weather will turn bad.

and considers repairing it by adding a causal premise:

- (C) The glass is falling.
Whenever the glass is falling the atmospheric pressure is falling.
Whenever the atmospheric pressure is falling the weather turns bad.
-

The weather will turn bad.

Now Railton considers (C) a causal explanation and its third premise a causal law. He points out that, until we understand the mechanism behind that causal relationship, we do not have an adequate explanation of the turn in the weather. He therefore concludes that we need something over and above the causal infusion that transformed (S) into (C). I agree that we need something more, but we also need something less.

Paul Humphreys has noted (in a personal communication) that there is something odd about the second premise of (C) as a component in a causal explanation, for its presence licenses an inference from effect to cause. And once that point is clear, we see that the first premise of that argument has no place in a causal explanation (or any other, I should think). As we noted in our earlier analysis of this example, the storm and the falling barometer are effects of a common cause, the drop in atmospheric pressure, and the cause screens the common effects off from one another. Therefore (C) should be replaced by

- (C') The atmospheric pressure is falling.
Whenever the atmospheric pressure is falling the weather turns bad.
-

The weather will turn bad.

The first premise of (C) is no part of the explanation; it is our evidence for the truth of the first premise of (C').

Still, I would not call (C') causal, for my conception of causality includes mechanisms of propagation and interaction. But I would regard the explanation that results from supplementation of (C')—in the way we both consider appropriate—a causal explanation. So far, then, the main disagreement is terminological. But Railton also makes mention of such structural laws as the Pauli exclusion principle, which he rightly holds to be noncausal (1978, 207). He characterizes explanations based upon such laws as noncausal *structural explanations*. In *Scientific Explanation and the Causal Structure of the World* I had taken the attitude that structural laws have explanatory force only if they themselves can be explained causally. In that context I was thinking of such structural laws as the ideal gas law, and I placed great weight upon its explanation on the basis of kinetic theory. Since I did not endeavor to offer an account of quantum mechanical explanation, I did not confront such laws as the Pauli exclusion principle.²⁴

At the time Railton sent this article off for publication he must have been working on his dissertation (1980). The brief suggestive remarks he offers in the article, in addition to his articulation of the D-N-P model, are elaborated clearly and at length in the dissertation—a monumental two-volume work that comes to 851 pages in toto.

When first I read the 1978 essay, I must confess, I failed to understand it. Only after seeing a later paper did I begin to appreciate the beauty of Railton's work on explanation. In 1981 a landmark symposium on probabilistic explanation (Fetzer 1981) was published; it included, among many other valuable contributions, Railton's "Probability, Explanation, and Information." In this essay he offers further discussion of the D-N-P model, addressing certain objections that might be brought against it. Of particular importance is the objection based on the acknowledged fact that many proffered explanations that are widely accepted as correct omit items, such as laws, that form an integral part of any explanation that conforms to the D-N-P model.

To deal with this problem, Railton introduces a distinction—one that turns out to be extraordinarily fruitful—between an *ideal explanatory text* and *explanatory information* (1981, 240). Given an event we wish to explain, the ideal explanatory text would spell out all of the causal and nomic connections that are relevant to its occurrence. For probabilistic explanation, the D-N-P model furnishes the schema for the ideal explanatory text. The ideal explanatory text can be expected—in most, if not all, cases—to be brutally large and complicated. When one considers the myriad molecules, atoms, subatomic particles, and interactions involved in everyday events, it is easy to see that the ideal explanatory text is an *ideal* that might never be realized. That does not matter.²⁵ The scientist, in seeking scientific understanding of aspects of our world, is searching for *explanatory information* that enables us to fill out parts of the ideal explanatory text. The ideal

explanatory text constitutes a framework that provides guidance for those who endeavor to achieve understanding of various aspects of the world, and that is its primary function.

Railton acknowledges the possibility that, in addition to the probabilistic explanations characterized by the D-N-P model, there may be nonprobabilistic explanations that are closely related to Hempel's D-N model. For Railton, however, D-N explanations are far more robust than they are for Hempel:

I would argue that the D-N schema instead provides the skeletal form for ideal explanatory texts of non-probabilistic phenomena, where these ideal texts in turn afford a yardstick against which to measure the explanatoriness of proffered explanations in precisely the same way that ideal D-N-P texts afford a yardstick for proffered explanations of chance phenomena. Thus, proffered explanations of non-probabilistic phenomena may take various forms and still be successful in virtue of communicating information about the relevant ideal text. For example, an ideal text for the explanation of the outcome of a causal process would look something like this: an inter-connected series of law-based accounts of all the nodes and links in the causal network culminating in the explanandum, complete with a fully detailed description of the causal mechanisms involved and theoretical derivations of all of the covering laws involved. This full-blown causal account would extend, via various relations of reduction and supervenience, to all levels of analysis, i.e., the ideal text would be closed under relations of causal dependence, reduction, and supervenience. It would be the whole story concerning why the explanandum occurred, relative to a correct theory of the lawful dependencies of the world. (1981, 246–47)

Does the conception of an ideal explanatory text have any utility? Railton replies,

[Is it] preposterous to suggest that any such ideal could exist for scientific explanation and understanding? Has anyone ever attempted or even wanted to construct an ideal causal or probabilistic text? It is not preposterous if we recognize that the actual ideal is not to *produce* such texts, but to have the ability (in principle) to produce arbitrary parts of them. It is thus irrelevant whether individual scientists ever set out to fill in ideal texts as wholes, since within the division of labor among scientists it is possible to find someone (or, more precisely, some group) interested in developing the ability to fill in virtually any aspect of ideal texts—macro or micro, fundamental or “phenomenological,” stretching over experimental or historical or geological or cosmological time. A chemist may be uninterested in how the reagents he handles came into being; a cosmologist may be interested in just that; a geologist may be interested in how those substances came to be distributed over the surface of the earth; an evolutionary biologist may be interested in how chemists (and the rest of us) came into being; an anthropologist or historian may be interested in how

man and material came into contact with one another. To the extent that there are links and nodes, at whatever level of analysis, which we could not even in principle fill in, we may say that we do not completely understand the phenomenon under study. (1981, 247–48)²⁶

The distinction between the ideal explanatory text and explanatory information can go a long way, I think, in reconciling the views of the pragmatists and the realists.²⁷ In his work on deductive-nomological explanation, Hempel steadfastly maintained an attitude of objectivity. D-N explanations were never epistemically relativized, and they always fulfill the requirement of maximal specificity trivially. As we have seen, Hempel found it necessary to relinquish this attitude when he confronted inductive-statistical explanations, but Coffa, Railton, and I, among others, made serious efforts to restore full objectivity to the domain of probabilistic explanation. At the same time, those philosophers who have emphasized the pragmatics of explanation, from Hanson and Scriven to van Fraassen, have criticized objectivist accounts for demanding inclusion of too much material to achieve what they regard as legitimate explanations. This opposition has, for example, led to the rejection of the covering law conception of explanation by pragmatists, and to insistence upon it by the objectivists.

The issue of how 'fat' or 'thin' explanations should be was raised in *Scientific Explanation and the Causal Structure of the World* (1984, 131–34), and it resurfaced in the 1985 American Philosophical Association Symposium on that book (Kitcher, van Fraassen, and W. Salmon, 1985). One useful way to think about this conflict, I believe, is to regard the objectivists—the advocates of the ontic conception—as focusing on the ideal explanatory text. We all hoped, I suspect, that the ideal explanatory text would be a good deal simpler than Railton had conceived it to be, but it did not work out that way. It was clear to me long ago, for example, in elaborating the statistical relevance model, that an objectively homogeneous partition, based on all objectively relevant factors, would usually have a horrendous number of cells, and, consequently, that a complete S-R explanation would be frightfully complex. However, if it is only the ideal text that is so forbidding, the situation is not so hopeless. The ideal explanatory text contains all of the objective aspects of the explanation; it is not affected by pragmatic considerations. It contains all *relevant* considerations.

When we turn to explanatory information, pragmatic considerations immediately loom large. What part of the ideal explanatory text should we try to illuminate? Whatever is *salient* in the context under consideration. That depends upon the interests and the background knowledge of whoever seeks the explanation. It depends upon the explanation-seeking why-question that is posed. Pragmatic considerations must not be seen as total determinants of what constitutes an adequate explanation, for whatever explanation is offered must contain explanatory information that coincides with something or other in the ideal explana-

tory text. The ideal explanatory text determines what constitutes explanatory information, and distinguishes it from explanatory misinformation. *Relevance* is a matter of objective fact; *salience* is a matter of personal or social interest. Thus, I should be inclined to say, the putative astrological explanation of President John F. Kennedy's death—discussed in §4.4—can be ruled out because it fails to coincide with any part of the ideal explanatory text. The configuration of stars and planets and satellites at the time of Kennedy's birth is (I firmly believe) irrelevant to the date of his assassination.

Looking at the pragmatics of explanation in this way, we can account for the rejections and asymmetries of explanation, which van Fraassen takes to be crucial (see Railton 1981, 248; fn 15 explicitly refers to van Fraassen). We can evaluate the presuppositions of why-questions to see if they are objectively satisfied. We can take necessary steps to determine—on the basis of contextual factors—just exactly what why-question is being posed. To do so we must specify van Fraassen's contrast class. We can consider the background knowledge of the individual who poses the question to ascertain what is missing—what knowledge gaps need to be filled if that person is to achieve scientific understanding—what aspect of the ideal text needs to be exhibited. We would also take into account the capacity of the questioner to assimilate scientific information in order to determine the depth and detail appropriate for that individual.

Given Railton's conception of explanatory information and the ideal explanatory text, it seems to me that much of the longstanding battle—going back to the early skirmishes between the logical empiricists and the ordinary language philosophers—can be resolved. I see this achievement as one foundation upon which a new consensus might be erected.

Railton refers to his general conception of scientific explanation as the *nomothetic account*; he compares it to the received view in the following way:

Where the orthodox covering-law account of explanation propounded by Hempel and others was right has been in claiming that explanatory practice in the sciences is in a central way *law-seeking* or *nomothetic*. Where it went wrong was in interpreting this fact as grounds for saying that any successful explanation must succeed either in virtue of explicitly invoking covering laws or by implicitly asserting the existence of such laws. It is difficult to dispute the claim that scientific explanatory practice—whether engaged in causal, probabilistic, reductive, or functional explanation—*aims* ultimately (though not exclusively) at uncovering laws. This aim is reflected in the account offered here in the structure of ideal explanatory texts: their backbone is a series of law-based deductions. But it is equally difficult to dispute the claim that many proffered explanations succeed in doing some genuine explaining *without* either using laws explicitly or (somehow) tacitly asserting their existence. This fact is reflected here in the analysis offered of explanatoriness, which is

treated as a matter of providing accurate information about the relevant ideal explanatory text, where this information may concern features of that text other than laws. (1981, 248–49)

In choosing to call his account *nomothetic* Railton emphasizes the role of laws. I think it equally deserves to be called *mechanistic*. We have already noted the crucial role of mechanisms in connection with the D-N-P model. In his 1981 article he remarks more generally on the standards for judging accounts of explanation:

The place to look for guidance is plainly scientific explanatory practice itself. If one inspects the best-developed explanations in physics or chemistry textbooks and monographs, one will observe that these accounts typically include not only derivations of lower-level laws and generalizations from higher-level theory and facts, but also attempts to *elucidate the mechanisms* at work. Thus an account of alpha-decay ordinarily does more than solve the wave-equation for given radionuclei and their alpha-particles; it also provides a model of the nucleus as a potential well, shows how alpha-decay is an example of the general phenomenon of potential-barrier penetration (“tunnelling”), discusses decay products and sequences, and so on. Some simplifying assumptions are invariably made, along with an expression of hope that as we learn more about the nucleus and the forces involved we will be able to give a more realistic physical model. It seems to me implausible to follow the old empiricist line and treat all these remarks on mechanisms, models, and so on as mere *marginalia*, incidental to the “real explanation”, the law-based inference to the explanandum. I do not have anything very definite to say about what would count as “elucidating the mechanisms at work”—probabilistic or otherwise—but it seems clear enough that an account of scientific explanation seeking fidelity to scientific explanatory practice should recognize that part of scientific ideals of explanation and understanding is a description of the mechanisms at work, where this includes, but is not merely, an invocation of the relevant laws. Theories broadly conceived, complete with fundamental notions about how nature works—corpuscularianism, action-at-a-distance theory, ether theory, atomic theory, elementary particle theory, the hoped-for unified field theory, etc.—not laws alone, are the touchstone in explanation. (1981, 242)

In view of these comments, and many others like them, I am inclined to consider Railton’s account primarily mechanistic and secondarily nomothetic.

Some time after I had seen his 1981 article, I happened to fall into conversation with Hempel—who was a colleague in Pittsburgh at the time—about Railton’s work. He mentioned that he had a copy of Railton’s dissertation (1980) and offered to lend it to me. After looking at parts of it rather quickly, I realized that I had found a treasure. Obtaining a copy of my own, I studied it at length, and

made considerable use of it in conducting the Minnesota Institute on Scientific Explanation in 1985. In a general survey article such as this I cannot attempt a full summary of its contents, but a little more should be said to indicate the scope of the theory he develops.

In his dissertation Railton distinguishes three kinds of explanations of particular facts. In response to various difficulties with D-N explanations of particular facts as treated by the received view, Railton maintains that an appeal to causality is required in many cases. He characterizes the *ideal causal nomothetic explanatory text for particular facts* and concludes that it "includes an account of the causal mechanisms involved at all levels, and of all the strands in the causal network that terminates in the explanandum" (1980, 725). However, he resists the notion that all particular-fact explanations are causal. There are, as we have seen, probabilistic explanations of particular facts, and we have discussed the appropriate type of ideal explanatory text. As an advocate of the propensity interpretation of probability, he regards probabilistic explanations as dispositional. But he is willing to take the concept of disposition in a sense broad enough to include dispositions of strengths zero and unity; consequently, he introduces, in connection with his second type of particular-fact explanation, the notion of the *ideal explanatory text for dispositional particular-fact explanations*, of which the D-N-P model provides a special case (1980, 735–36). The third type of particular-fact explanation, for Railton, is structural. Such explanations appeal typically to such laws of coexistence as the Pauli exclusion principle, Archimedes' principle, laws of conservation, and so on. Maintaining that such explanations cannot plausibly be construed as disguised causal explanations, he also introduces the notion of the *ideal explanatory text for structural particular-fact explanations* (1980, 736–39). Explanations of this type—like causal and dispositional explanations—involve essential reference to the underlying mechanisms.

Recognizing that actual explanations of particular occurrences may involve elements of more than one of the foregoing, Railton also gives us the notion of a *nomothetic ideal encyclopedic text* (1980, 739). "Harmonious co-operation among the elements of an encyclopedic text is possible because all three forms reflect a conception of explanation that we may intuitively describe as this: an explanation should invoke the factors, laws, etc., that actually *bring about or are responsible for* the explanandum (once these relations are broadened beyond the purely causal), i.e., the explanations should show what features of the world the explanandum is *due to*" (1980, 739–40).

Because of certain differences in terminology among the three of us, perhaps it would be helpful to say something about the relationships among Coffa, Railton, and me. First, it should be noted, I am an advocate of probabilistic causality, so when I speak of causal explanation it is explicitly intended to include probabilistic explanation. Both Coffa and Railton use the term "cause" in the narrower deterministic sense. Furthermore, Coffa and Railton advocate the propen-

sity interpretation of probability, while I reject the notion that propensities satisfy the axioms of probability, and hence that there is any such thing as a propensity interpretation. However, I do believe that the concept of a probabilistic propensity is an extremely useful notion; I would identify it with probabilistic cause. Hence, not to put too fine a point on it, Coffa, Railton, and I all agree on the general idea of a causal or dispositional type of explanation that coincides with Railton's causal and dispositional types. Up to this point, the differences are mainly terminological. A major difference arises, however, between Coffa on the one hand and Railton and me on the other, for both of us place far more importance upon an appeal to mechanisms than does Coffa. On another point Railton differs from Coffa and me; neither of us accords a distinctive place to Railton's structural explanations. My *constitutive explanations* bear some resemblance, but I regard them as fundamentally causal (1984, 270–76), whereas Railton explicitly denies that his structural explanations are. While I am still inclined to think that many of Railton's structural explanations can be analyzed causally, I am far from confident that all can. That point seems particularly clear with regard to such quantum mechanical examples as the Pauli exclusion principle, but, as we know, the nature of quantum mechanical explanation is deeply perplexing (W. Salmon 1984, 242–59). Nevertheless, had I read Railton's dissertation before the publication of my book, I would have devoted considerably more attention to structural explanation.

Railton's overall account is not confined to particular-fact explanations; he also offers an account of theoretical explanation. In summarizing his approach he remarks,

It was argued that by beginning with particular-fact explanation, we had not prejudiced the account against theoretical explanation, for, in fact, theoretical explanation had been involved all along in the full development of ideal explanatory texts. As might have been expected, the nomothetic account recognizes that regularities and laws may have causal, dispositional, or structural elements in their explanations, and these elements have virtually the same ideal forms as in particular-fact explanation. Thus the same requirements of true ingredients, basic covering-law structure, an account of mechanisms, thorough-going theoretical derivation, asymmetries, etc., apply in theoretical explanation. The nomothetic account thereby preserves the estimable unity of theoretical and particular-fact explanation that is characteristic of covering-law approaches to explanation. (1981, 746–47)²⁸

Railton's dissertation offers a deep and sensitive treatment of a wide range of problems, issues, and views regarding the nature of scientific explanation. Although it does not attempt to resolve some of the most fundamental problems we have encountered in our discussions of explanations – such problems as the nature of laws, the analysis of causality, the nature of mechanisms, the notion of a purely

qualitative predicate—it is a rich source of philosophical insight on the nature of scientific explanation. While I do not by any means agree with all of his views, I do believe that anyone who is seriously interested in philosophical work on scientific explanation should study his dissertation with care.

4.7 Aleatory Explanation: Statistical vs. Causal Relevance

In the early part of the third decade—when I was busily expounding the statistical-relevance model—I was aware that explanation involves causality, but I hoped that the required causal relations could be fully explicated by means of such statistical concepts as screening off and the conjunctive fork. A decade later, I was quite thoroughly convinced that this hope could not be fulfilled (W. Salmon 1980; see also 1984, chap. 7). Along with this realization came the recognition that statistical relevance relations, in and of themselves, have no explanatory force. They have significance for scientific explanation only insofar as they provide evidence for causal relations. By 1984 (34–47) they had been relegated to *the S-R basis* upon which causal explanations can be founded. Causal explanation, I argued, must appeal to such mechanisms as causal propagation and causal interactions, which are not explicated in statistical terms.

The question arises of what to do with the S-R basis. Given the fact that it will often be dreadfully complex, we might consign it to Railton's ideal explanatory text, recognizing that it is the sort of thing that will not often be spelled out explicitly. We will refer to parts of it when we need to substantiate causal claims.

Another sensible approach to this problem is, in the words of Frank Lloyd Wright, "Abandon it!"²⁹ This is the tack taken by Paul Humphreys in articulating his theory of *aleatory explanation* (1981, 1983). A basic difference between Humphreys's model and other models of probabilistic or statistical explanation extant at the time is that all of the latter require one or more probability values to appear explicitly in the completed explanation. Although knowledge of probabilities is used in constructing aleatory explanations, values of probabilities are absent from the explanation itself.

According to Humphreys, factors that are causally relevant are also statistically relevant. Just as there are two kinds of statistical relevance—positive and negative—so also are there two kinds of causes. Causes that tend to bring about a given effect are *contributing causes*; those that tend to prevent the effect are *counteracting causes*. The canonical form for an aleatory explanation is "A because Φ , despite Ψ ", where Φ is a non-empty set of contributing causes, Ψ is a set, possibly empty, of counteracting causes, and A is a sentence describing what is to be explained." It is assumed that Φ and Ψ include all known causally relevant factors. The set of contributing causes must not be empty, for we have no explanation at all if only counteracting causes are present.

Consider a modified version of one of Humphreys's examples. Suppose that

a car has gone off a concrete road at a curve and that the only known conditions that are causally relevant are the fact that the driver was fully alert and the car was traveling at an excessive speed. The first is obviously a counteracting cause of the accident; the latter a contributing cause. We might then say that the car went off the road because it was traveling too fast, despite the fact that the driver was alert. There are, of course, many other statistically relevant factors, and if they are taken into account the probability of the explanandum will change. But as long as there are no other factors that screen off excessive speed, or render it a counteracting cause, its status as a contributing cause is unchanged. Similarly, if there are no other factors that screen off driver alertness or transform it into a contributing cause, its status as a counteracting cause holds.

Suppose now we find out that, in addition, visibility was clear, but there was sand on the road at the curve. The first of these is a counteracting cause of the accident, and the second is a contributing cause. We may add both to the original explanation, transforming it into the following: The car went off the road, despite the fact that the driver was alert and visibility was clear, because the car was traveling too fast and there was sand on the road at the curve. The result is that the first explanation, though incomplete, is correct; we have not been forced to retract any part of the original. If, in contrast, we had been trying to construct an S-R explanation of the same phenomenon, we would have been required to retract the original statistical relevance relations and replace them with others.

In constructing aleatory explanations, we must be aware of the possibility that the introduction of an additional causally relevant factor may change a contributing cause into a counteracting cause or vice-versa, or it may render a cause of either type irrelevant. Suppose we learn that there was ice on the road at this curve. This would change the presence of sand on the road from a contributing cause to a counteracting cause, for if the road were icy the sand would tend to prevent the car from skidding. In this case, the additional factor would force us to reject the former explanation, even as a correct partial explanation.

When constructing aleatory explanations, we are aware of the danger of a defeating condition—i.e., a condition, such as ice on the road in the preceding example, that transforms a contributing cause into a counteracting cause or vice-versa. By careful investigation we try to assure ourselves that none are present. Establishing such a result would be analogous to verifying one of Coffa's extremal clauses (§3.3). Just as we might check for the absence of compressing forces before applying the law of thermal expansion to a heated iron bar, so also would we check for road surface conditions that could transform sand on the road from a contributing to a counteracting cause before offering an explanation of the car leaving the road.

Humphreys's theory of aleatory explanation falls clearly within the ontic conception; it constitutes a valuable contribution to our understanding of causal explanation, where probabilistic causes are included. Inasmuch as its fullest articu-

lation is contained in "Scientific Explanation: The Causes, Some of the Causes, and Nothing but the Causes," his contribution to this volume (Kitcher and Salmon 1989), I shall resist the temptation to discuss his theory at greater length and let him speak for himself.

4.8 Probabilistic Causality

In the third decade, as we saw, Coffa's theory placed the notion of a probabilistic disposition (what I would call a probabilistic cause) in a central position with respect to scientific explanation. During the same decade, a similar line of thought was pursued by Fetzer, whose theory was fully articulated early in the fourth decade. By the beginning of the fourth decade I was thinking seriously about how the S-R (statistical-relevance) model could be augmented by suitable causal considerations. At that time I was aware of only three theories of probabilistic causality that had been reasonably well worked out, namely, those of Reichenbach (1956), I. J. Good (1961–62), and Patrick Suppes (1970). Neither Reichenbach nor Suppes drew any connection between his theory of probabilistic causality and scientific explanation; Good used his mathematical definition of degree of explicativity (1977) in a key role in his probabilistic causal calculus, but without much philosophical elaboration. I surveyed these theories (1980) and pointed to severe difficulties in each; soon thereafter Richard Otte (1981) exhibited still greater problems in Suppes's theory. The primary moral I drew was that causal concepts cannot be fully explicated in terms of statistical relationships; in addition, I concluded, we need to appeal to causal processes and causal interactions. (For details see W. Salmon 1984, chaps. 5–7.)

Early in the fourth decade Fetzer and Nute (1979) published a new theory of probabilistic causality that was intended to play a crucial role in the explication of scientific explanation.³⁰ Fetzer's theory of scientific explanation is based on two cardinal principles: first, the interpretation of probability as a single-case propensity, and second, the inadequacy of extensional logic for the explication of such fundamental concepts as lawlikeness and causality. Single-case propensities are understood as probabilistic dispositions, indeed, as probabilistic causes. To elaborate a theory of scientific explanation that embodies these ideas, Fetzer and Nute construct a modal logic in which conditional statements embodying three kinds of special connectives are introduced, namely, subjunctive conditionals, causal conditionals (involving dispositions of universal strength), and probabilistic conditionals (involving dispositions having numerical degrees of strength). Those of the last type are statements of probabilistic causality.

The formal system they construct has 26 axiom schemas in all. The first six form the basis for the logic of subjunctive conditionals; the symbol for the connective is a fork ($\supset$). This logic is not different in kind from various well-known systems of modal logic. The next eight axiom schemas form the basis for the logic

of universal causal conditionals; the symbol for this connective is the u-fork ($\exists_u$). There is nothing wildly nonstandard in this part of the system either. The final dozen axiom schemas form the basis for the logic of probabilistic causality. At this stage, however, instead of introducing one new connective symbol, the authors introduce a *nondenumerable infinity* of symbols ($\exists_n$), the n-forks, where n assumes the value of each real number in the unit interval. They then talk about establishing the well-formed formulas of this calculus "in the usual way."

What they have done is, on the contrary, most unusual. In the first place, the standard representation of real numbers is by means of sequences of digits, almost all of which are infinite sequences. It appears, then, that the vast majority of symbols of the form " $\exists_n$ " are not single symbols, or finite strings of symbols, but infinitely long strings. Standard logical and mathematical languages, though they usually admit countable infinities of symbols, generally limit the well-formed formulas to finite length. In standard logical and mathematical languages there can exist names for no more than a denumerable subset of the real numbers. The use of a nondenumerable infinity of symbols and infinitely long well-formed formulas signal fundamental difficulties for the Fetzer-Nute proposed system.

In several of the axiom schemas, for example, the n-fork appears twice with the same subscript. If we look at a given formula containing two n-forks with their numerical subscripts—one that coincides with an axiom schema in all other respects than the equality or inequality of the two subscripts—we must determine whether the subscripts are identical to ascertain whether the formula is an axiom. This is done by comparing the two sequences digit by digit. If the two sequences are different, we will discover a discrepancy within a finite number of comparisons; but if they are the same we will never be able to establish that fact. There is no effective way of deciding, in general, whether two representations designate the same real number.

One of the basic virtues of a good axiomatic system is that it provides us with the capability of recognizing an axiom when we see one. If we have a finite number of finite formulas that capability is evident. In addition, given a finite number of *suitable* axiom schemas, each of which admits an infinite set of axioms, it is possible to provide an effective method for recognizing an axiom when we meet one. The Fetzer-Nute system, as presented, does not give us any such method. Unless the formation rules are spelled out in full detail, and a way is found to circumvent this problem, the probabilistic part of the calculus cannot get off the ground. Fetzer and Nute acknowledge that their nondenumerable infinity of well-formed formulas presents a difficulty when it comes to an attempt to establish the completeness of this part of their calculus (Fetzer and Nute 1979, 473; Fetzer 1981, 67), but I think the problem lies much deeper.

Another conspicuous feature of the probabilistic part of the calculus is that from twelve rather complicated axioms only four utterly trivial theorems are offered:

If p probabilistically implies q , then

- (a) p does not necessarily imply q , and p does not necessarily imply not- q ;
- (b) p does not have a universal disposition to produce q , and p does not have a universal disposition to produce not- q ;³¹
- (c) p is possible and not- q is possible;³²
- (d) p and q are jointly possible.

Such a large tree should bear more fruit.

The foregoing criticisms are not meant to imply that all attempts to construct a probabilistic causal calculus within a nonextensional logic are bound to be futile or fruitless, but only that the Fetzer-Nute version needs further work if it is to succeed.

In the less formal portion of his work on scientific explanation, Fetzer offers searching discussions of Hempel's requirement of maximal specificity and of the homogeneity requirement I imposed on the S-R model. As we recall, in 1968 Hempel replaced his 1965 RMS with a revised version RMS*. Fetzer offers further revisions and offers RMS** which, he believes, deals adequately with the problems concerning relevance that I had raised (1981, 99). In addition, he offers a revised explication of homogeneity designed to escape certain difficulties he alleged to be present in mine.

Two main points of his critique are worth at least brief comment. The most serious, I think, involves "the mistaken identification of statistical relevance with explanatory relevance" (1981, 93), since "*statistically relevant* properties are not necessarily *causally relevant* (or *nomically relevant*) properties, and conversely . . . " (1981, 92). This is a criticism whose validity I have completely endorsed, as I have remarked repeatedly in foregoing sections. When the S-R model was first published I believed that causal relevance could be explicated entirely in terms of statistical relevance relations, making great use of the screening off relation. In that place, however, I did not attempt to carry out any such analysis. During the next several years I became increasingly skeptical about the viability of any such approach, arriving finally at the fairly strong conviction that it could not be done (1980). However, the details of the mechanisms of causality and their role in scientific explanation were not spelled out in full detail until (1984), three years after the publication of Fetzer's book.

Another one of Fetzer's basic arguments is that, according to my definition, the only objectively homogeneous reference classes are those that have only one member, and, consequently, the relative frequency of any attribute in such a class is, of logical necessity, either zero or one (1981, 86-94). Therefore, he claims, whereas I had accused Hempel's theory of an implicit commitment to determinism, in fact my account is at least equally guilty. As I had long been aware, the reference class problem is extremely serious for any theory of statistical explanation. In Hempel's account, RMS (later RMS*) was designed to deal with this

problem; in mine, objective homogeneity was intended to handle it. In my initial presentations of the S-R model, early in the third decade, I handled the problem quite cavalierly, making passing reference to Richard von Mises's concept of a place selection. By the end of the third decade, I realized that it had to be taken very seriously. My first attempt at a detailed explication of objective homogeneity was badly flawed (1977b); an improved treatment, which I hope is more successful, was not published until 1984 (chap. 3). However that may be, I am still convinced that the concept of an *objectively homogeneous reference class* is legitimate and important. I maintain, for example, that the class of carbon-14 atoms is objectively homogeneous with respect to the attribute of spontaneous radioactive decay within 5730 years. If our concept does not fit cases of that sort, our explication must be at fault.

Obviously I cannot accept Fetzer's view that problems with objective homogeneity are merely unfortunate consequences of adopting of the *extensional* frequency interpretation of probability; they are problems that can be solved and deserve to be solved. Moreover, as I have argued in detail (1979a), the single-case propensity interpretation does not escape what amounts to the same type of problem, namely, the specification of the chance set-up that is supposed to possess the probabilistic disposition, for on any propensity interpretation it is necessary to specify what counts as repeating the experiment. Fetzer addresses this problem by imposing *the requirement of strict maximal specificity*:

An explanation of why an explanandum event . . . occurs is adequate *only if* every property described by the antecedent condition(s) is nomically relevant to the occurrence of its attribute property . . . (1981, 125–26)

Nomic relevance is relevance by virtue of a causal or noncausal law. Nomic relevance is emphatically not to be identified with statistical relevance.

Fetzer uses this requirement in the formulation of his characterization of causal explanation:

A set of sentences S, known as the "explanans," provides *an adequate nomically significant causal explanation of the occurrence of a singular event described by another sentence E, known as its explanandum, relative to [a given] language framework*, if and only if:

- (a) the explanandum is either a deductive or a probabilistic consequence of its explanans;
- (b) the explanans contains at least one lawlike sentence (of universal or statistical) 'causal' form that is actually required for the deduction or probabilistic derivation of the explanandum from its explanans;
- (c) the explanans satisfies the requirement of strict maximal specificity (RSMs) with respect to its lawlike premise(s); and

- (d) the sentences constituting the explanation—both the explanans and the explanandum—are true, relative to the [given] language framework. (1981, 126–27)

As a gloss on this formulation, Fetzer remarks, “If the law(s) invoked in the explanans are essentially universal, the logical properties of the relationship between the sentences constituting the explanans and its explanandum will be those of complete (deductive) entailment; while if they are essentially statistical, this relationship will be that of only partial (deductive) entailment. The logical relation, in either case, is *strictly deductive*” (1981, 127). In the following chapter, Fetzer offers an entirely parallel analysis of *nomically significant theoretical explanation*, a noncausal form of explanation.

Fetzer has offered a nonextensional explication which, by virtue of its appeal to partial entailment, bears striking resemblance to Mellor’s nondeterministic version of the modal conception of scientific explanation. In commenting (above) on Mellor’s views I expressed my strong doubts about the viability of the concept of partial entailment; these qualms apply equally to Fetzer’s employment of it. Given the intensional analysis, however, Fetzer has a straightforward answer to the issue we raised in §4.3 concerning the relationship between descriptive and explanatory knowledge. He suggests that the distinction between *description and prediction*, on the one hand, and *explanation*, on the other, is that the former can proceed in an extensional language framework, while the latter demands an intensional language framework. It remains to be seen whether the intensional logic can be satisfactorily formulated.

Fetzer has offered an account of explanation that is, in an extended sense, deductive. His book is dedicated to Karl Popper. In the next section we shall consider a more orthodox Popperian approach to statistical explanation.

4.9 Deductivism

The thesis that all legitimate scientific explanations are deductive arguments has a long and proud history, going back at least to Aristotle. It has been reiterated by many philosophers, including John Stuart Mill in the nineteenth century, and Karl R. Popper in the early part of the twentieth century. During our four decades it has been advocated by Brodbeck, Stegmüller, and von Wright, as well as many others, including Popper and his followers. It is *not*, however, a view that Hempel ever explicitly held, for he steadfastly maintained that there are, in addition, explanations of the inductive-statistical type.³³

Before the twentieth century, during the reign of classical physics and Laplacean determinism, deductivism with respect to scientific explanation was a natural and appealing view. One might have tolerated something like explanations of the inductive-statistical variety as long as it was clearly understood that they were in-

complete, the appeal to probability or induction being merely a result of our ignorance of the full explanatory laws or facts. With the advent of quantum mechanics it became necessary to admit that determinism might well be false, and that deductive-nomological explanations of some important phenomena may be impossible in principle.

One possible response to these new developments in physics is simply to hold stubbornly to determinism; this reaction seems utterly anachronistic. New—as yet undiscovered—physical theories may eventually convince us that determinism is true; even so, we have no *a priori* guarantee that determinism will be reinstated. At this point in the history of physics our philosophical theories of explanation must leave open the possibility that the world is indeterministic.

Another possible response is to admit that quantum mechanics is an indeterministic theory, but then to deny that quantum mechanics furnishes explanations of any physical phenomena. This reaction also seems unwarranted. Quantum mechanics (including quantum electrodynamics and quantum chromodynamics) has had more explanatory success than any other theory in the history of science. Classical physics could not explain the distribution of energy in the spectrum of blackbody radiation; quantum mechanics provided the explanation. Classical physics could not explain the photoelectric effect; quantum mechanics could. Classical physics could not explain the stability of atoms; quantum mechanics could. Classical physics could not explain the discrete spectrum of hydrogen; quantum mechanics could. Unless one wants to retreat to the old position that science never explains anything, it seems implausible in the extreme to deny that quantum mechanics has enormous explanatory power.

Another avenue is, however, open to the deductivist. Without relinquishing the deductivist position, one can concede that quantum mechanics has great explanatory power, but that all of its explanations are of the type Hempel characterized as deductive-statistical. This means, of course, that quantum mechanics can furnish no explanations of particular facts; the only things it can explain are statistical laws. To take just one example, the deductivist who adopts this stance must maintain that, while quantum mechanics can explain how, in general, electrons are scattered by crystals, it cannot explain the particular patterns actually obtained by Davisson and Germer in their famous experiment.³⁴ I find it difficult to accept this conclusion. Given the fact that quantum mechanics governs the microstructure of all macroscopic phenomena, it would appear to lead to the conclusion that science cannot explain any particular occurrences whatever. As I argue in a paper that will appear in the first year of the fifth decade (1988), even if theoretical science could be shown to have no need of explanations of particular facts, it seems impossible to make the same claim for applied science. When, for example, questions of legal liability arise, we seek causal explanations (which may have probabilistic components) for such particular occurrences as the collapse of a tank holding vast quantities of oil or the contracting of lung cancer by

a long-term heavy smoker of cigarettes. Philosophy of science that confines its attention to pure science, to the complete neglect of applied science, is, I suggest, severely biased.

Another maneuver available to the deductivist is to maintain that, although we cannot have a scientific explanation of a particular chance event, we can have an explanation of the fact that such an event has some particular probability. An easy way to see how it would work is to reconsider Railton's D-N-P model (discussed in section 4.6). In focusing on Railton's model, it is essential to keep in mind that he is not a deductivist. According to Railton, we can have explanations of chance events; such explanations consist of two parts: (1) a deductive argument, whose conclusion is that the event in question has a certain probability, and (2) a parenthetical addendum, which states that the event in fact occurred. If we were to accept the first part, while rejecting the second, we would be left with a deductive explanation of the fact that the explanandum-event has a particular probability.³⁵ Such a view leaves us in the position of having to say that science cannot explain what happens in the world; it can only explain why those things that do happen have certain probabilities. It can also explain why things that do not happen have certain probabilities of occurring.

When Richard Jeffrey challenged Hempel's claim that statistical explanations are arguments, he excepted certain "beautiful cases" in which the probability of occurrence is so great that there seems no point in giving any weight at all to their non-occurrence. One of his examples is the failure of a flat tire to reinflate spontaneously as a result of a jet of air formed by chance from the random motion of molecules in the surrounding air. Another example would be the melting of an ice-cube placed in a glass of tepid water. Although he allowed that such explanations may be construed as arguments, he did not, of course, claim that they are deductive. Railton maintained, on the contrary, that the mere fact that some probabilities are high and some not so high does not make a difference in principle between the "beautiful" and less attractive cases. If the latter are not arguments then neither are the former. My own view is that Railton is correct in this observation, but my opposition to the inferential conception of explanation is so deep that I may simply be prejudiced.

One staunch deductivist who has recently considered the problem of Jeffrey's "beautiful cases" is John Watkins, whose neo-Popperian *Science and Scepticism* appeared in 1984. Although he is a deductivist, he maintains that, while science cannot provide explanations of individual micro-events, it can furnish explanations of macro-events that consist of large aggregates of chance micro-events:

There is a far-reaching analogy between the explanatory power of the deterministic theories of classical physics and the indeterministic theories of modern micro-physics . . . Both explain empirical regularities by appealing to higher level structural laws that are taken as immutable and absolute. In the

case of classical physics, such a law says that, given only that such-and-such conditions are satisfied, nothing whatever can prevent a certain outcome from following. In the case of microphysics, it says that, given only that such-and-such conditions are satisfied, nothing whatever can alter *the chance* that a certain outcome will follow. One could as well call the latter an "iron law" of chance as the former an "iron law" of nomic necessity. Both kinds of law, in conjunction with appropriate initial conditions, can explain empirical regularities and, indeed, singular macro-events (provided that the macro-event in question is the resultant of a huge aggregate of micro-events . . .). The analogy breaks down, however, when we come down to individual events at the micro-level. . . . I have argued that it is a mistake to call upon microphysics to *explain* an individual micro-event, such as the disintegration of a radon atom or the reflection of a photon; for if it really was a matter of chance which way it went, then the fact that it chanced to go this way rather than that simply defies explanation. What we *can* explain with the help of an appropriate indeterministic microphysical theory is why there was a precise objective probability that it would go this way. (1984, 246)

To support his claim that particular macro-events can be explained, Watkins discusses one of Hempel's familiar examples. Suppose we have a sample consisting of 10 milligrams of radon. We find, after 7.64 days (two half-lives of radon), that the sample contains 2.5 ± 0.1 milligrams of radon. Given the law of spontaneous radioactive decay and the fact that the original sample contained more than 10^{19} atoms, we can deduce the probability that at the end of 7.64 days 2.5 ± 0.1 milligrams of radon will remain. This probability is extremely close to unity; indeed, it differs from one by less than $10^{-10^{15}}$

At this point, anyone, whether a deductivist, inductivist, or whatever, concerned with the nature of microphysics, faces a crucial question: what *physical* meaning should be given to vanishingly small values [of probability]? Should we interpret our disintegration law as allowing that the macro-outcome *might* fall outside the [given] interval? Should we interpret the laws of thermodynamics as allowing that patches of ice and wisps of steam *might* form spontaneously in an ordinary bathtub because of hugely improbable distributions of molecules? (1984, 243)

Following an approach to statistical laws that had been advocated by Popper (1959, 198–205), Watkins argues that in cases like Hempel's radon decay example, the statistical law is *physically equivalent* to a universal law, namely, given *any* 10 mg. sample of radon, it will contain 2.5 ± 0.1 mg. of radon after two half lives have transpired. If the statistical decay law is replaced by its universal surrogate, we can, of course, furnish a deductive explanation of the phenomenon. This is the foundation of Watkins's claim that, although we cannot explain in-

dividual micro-occurrences, we can provide explanations of macro-events that involve the behavior of extremely large numbers of micro-entities.³⁶

Whether one accepts or rejects the Popper-Watkins thesis about deductive explanations of large aggregates of micro-events, the deductivist position—which rejects all such models of explanation as the I-S or S-R—is appealing because of its avoidance of the problems associated with maximal specificity, epistemic ambiguity, and explanations of improbable events. According to deductivism, given an indeterministic world, we must forego explanations of *what actually happens* (in many cases, at least); however, we can, by revealing the stochastic mechanisms, understand how the world works. I shall not go into greater detail about the merits of deductivism here, since Philip Kitcher's contribution to this volume (Kitcher & Salmon 1989) contains an extensive elaboration of that position.

There is, however, a residual problem with deductivism that merits attention. It has to do with the relationship of causality to explanation. In his celebrated work *The Cement of the Universe*, J. L. Mackie invites consideration of three machines that dispense candy bars (1974, 40–43). One of the machines is deterministic; the other two are indeterministic. Leaving aside some details that are inessential to our discussion, we can say that the deterministic machine gives you a candy bar if and only if you insert a shilling.

Among the indeterministic machines, the first gives a candy bar only if a shilling is inserted, but it sometimes fails to give one when the coin is put in. There is no deterministic explanation of these failures; they simply happen occasionally by chance. Suppose a shilling is inserted and the machine yields a candy bar. In this case, according to Mackie, putting the coin in the slot *causes* the candy bar to come out, for without the coin there would have been no candy bar. This is the *sine qua non* conception of causality; a cause is a necessary condition. I agree with Mackie about the causal relation here, and I would add that putting the coin in the slot *explains* the ejection of the candy bar.

The second indeterministic machine is the converse of the first. Whenever a shilling is inserted a candy bar is forthcoming, but occasionally, by chance, the machine ejects a candy bar when no coin is put in. The insertion of the coin is a sufficient, but not a necessary, condition of getting a candy bar. Suppose someone puts a shilling in this machine and receives a candy bar. According to Mackie it would be wrong to say that putting in the coin causes the candy bar to come out, for a candy bar might have been forthcoming even if no coin had been inserted. Again, I think that Mackie is right, and I would go on to say that the insertion of the coin does *not* furnish a (nonstatistical) explanation of the appearance of the candy bar.

If the intuition about these cases—shared by Mackie and me—is correct, it leaves the deductivist in an awkward position regarding scientific explanation. In the case of the first machine, we have identified the cause of the explanandum-event, but we cannot provide a D-N explanation, for the insertion of the coin is

not a sufficient condition of getting a candy bar. In the case of the second machine, we can provide a D-N *explanation* of the explanandum-event, but there is no (nonprobabilistic) cause of it.

In my discussion of this point, to escape the blatant artificiality of Mackie's candy machines, I introduced two photon-detectors. The first never produces a click unless a photon impinges on it, but it fails to click in a small percentage of cases in which a photon impinges. The second detector never fails to click when a photon impinges, but occasionally it gives a spurious click when no photon is present. These detectors are strictly analogous to the candy-dispensing machines. With the first detector we have the cause of a click, but (according to the deductivist) no explanation. With the second detector we have (according to the deductivist) an explanation, but no cause. The deductivist, it seems to me, needs to come to terms with examples of this sort (W. Salmon 1988).

4.10 Explanations of Laws Again

There is one further point in Watkins's book that has important bearing on one of the recalcitrant problems we have encountered regarding deductive explanation. It will be recalled that Hempel and Oppenheim—in their notorious footnote 33—explained why they offered no account of deductive explanation of laws. Their difficulty was that Kepler's laws *K* could be deduced from the conjunction of those very laws with Boyle's law *B*, but this would surely fail to qualify as a bona fide explanation. The problem is to characterize precisely the distinction between deductions of that sort, which do not constitute explanations, and those derivations of regularities from more general laws that do constitute legitimate explanations.

According to Popperians, the fundamental aim of science is to produce and test bold explanatory theories. Watkins regards scientific theories as finite sets of axioms. The axioms must be logically compatible with each other and they must be mutually independent. He remarks, "It is rather remarkable that, although scientific theories are taken as the basic units by many philosophies and nearly all histories of science, there is no extant criterion, so far as I am aware, for distinguishing between a theory and an assemblage of propositions which, while it may have much testable content, remains a rag-bag collection" (1984, 204). As an answer to this problem he offers what he calls *the organic fertility requirement* (1984, 205). If a theory *T* contains more than one axiom, then it fulfills this requirement if it is impossible to partition the axiom set into two mutually exclusive and exhaustive nonempty subsets *T'* and *T''*, such that the testable content of *T* is equal to the sum of the testable contents of *T'* and *T''*. In other words, the axioms must work together to yield testable consequences that they cannot generate separately. With reasonable restrictions on what qualifies as testable content, it seems clear that the testable content of Kepler's laws and Boyle's law is no greater than the

set theoretical union of the testable content of Kepler's laws and the testable content of Boyle's law. Kepler's three laws — K_1 , K_2 , K_3 — presumably satisfy the organic fertility requirement, while the set consisting of those together with B would not.

The situation is somewhat complicated by the fact that any given theory can be axiomatized in many different ways. We could, for instance, form the conjunction of all four laws to give us a theory with just one axiom. To deal with this kind of move, Watkins provides five rules for "natural" axiom sets. Among them is one he designates as "Wajsberg's requirement"; it says, in part, "An axiom is impermissible if it contains a (proper) component that is a theorem of the axiom set" (1984, 208). This clearly disposes of the Hempel-Oppenheim example. Whether Watkins's requirements block all counterexamples is a question I shall not try to answer. If they do, then they can be used to solve the problem with which Michael Friedman was concerned. Theories that satisfy the organic fertility requirement and are axiomatized "naturally" serve to unify our scientific knowledge.

4.11 A Fundamental Principle Challenged

There is a principle that has long been considered a cornerstone in the theory of scientific explanation, namely, if a set of circumstances of type C on one occasion explains the occurrence of an event of type E, then circumstances of the same type C cannot, on another occasion, explain the nonoccurrence of an event of type E (or the occurrence of an event of a type E' that is incompatible with E). Since it has so often been taken as a first principle, let us call it "Principle I." If this principle were relinquished, it has been thought, the floodgates would be open to all sorts of pseudo-explanations that are scientifically unacceptable. Nevertheless, careful consideration of probabilistic or statistical explanation has led some authors to reject that principle.

The D-N model of scientific explanation clearly satisfies this principle, for it is impossible validly to deduce two incompatible propositions from any consistent set of premises. Hempel's I-S model also satisfies it, as long as the high-probability requirement is enforced; since the sum of the probabilities of two incompatible statements with respect to any given consistent body of evidence cannot exceed one, it is impossible for both to have high probabilities. Adherents of the modal conception are committed to Principle I, for to show that an occurrence is necessary obviously implies that any incompatible occurrence is impossible. In van Fraassen's pragmatic theory Principle I is embodied in the view that an explanation shows why the topic *rather than* any other member of the contrast class is true.

At various junctures in our discussion of statistical explanation, I have made reference to a symmetry principle, namely, if a given stochastic process gives rise to some outcomes that are highly probable and to others that are improbable, then we understand the improbable ones just as well (or as poorly) as the probable

ones. If this symmetry principle is correct—and I believe it is—it puts us in a dilemma. It forces us to choose between abandoning Principle I or forgoing explanations of probabilistic outcomes. For if circumstances C explain the probable outcome E on many occasions, then the same circumstances C explain the improbable outcome E' on some other occasions. For closely related reasons, both Stegmüller and von Wright rejected the claim that probabilistic explanations of particular occurrences are possible. Since I am not inclined to give up statistical explanations of single events, and since I am strongly opposed to any high probability requirement, I have (in company with Achinstein and Railton) rejected Principle I. Although there may be non-deductivists—such as Hempel in “Aspects of Scientific Explanation”—who reject the symmetry principle, thus embracing Principle I and statistical explanations of particulars, Principle I does seem to be one of the chief bludgeons of the deductivists (see, e.g., Watkins 1984, 246).

What are the hazards involved in the rejection of Principle I? We have often been warned—and rightly so—that science has no place for theological or metaphysical ‘theories’ that explain whatever happens. Suppose someone is critically ill. If the person dies, the loved ones explain it as owing to “God’s will.” If the person recovers, they explain it as owing to “God’s will.” Whatever happens is explained in terms of the will of the Almighty. However comforting such ‘explanations’ might be, they are vacuous because there is no independent way of determining just what God wills. To rule out ‘explanations’ of this sort we do not need to appeal to Principle I; it is sufficient to insist that scientific explanations invoke *scientific* laws and facts. Scientific assertions, including those employed for purposes of explanation, should be supported by evidence. If we are dealing with what Hempel and Oppenheim called “potential explanation,” then the laws or theories involved must be capable of independent support.

At the outset of our story we mentioned the attitude of scientific philosophers toward Driesch’s attempts to explain biological phenomena in terms of entelechies and vital forces. During the same era, Popper was severely critical of Freudian psychoanalytic theory and Marxian economics. In all of these cases the criticisms were directed against the empirical vacuousness of the theories involved. To the extent that the theories in question are, indeed, empirically vacuous, to that extent they *are* devoid of scientific explanatory import. Such theories differ radically from the basic statistical theories of contemporary physics. These statistical theories offer a range of possible outcomes from a given indeterministic situation, attaching a definite probability value to each. They are far from vacuous, and they are supported by a vast amount of empirical evidence. The fact that such theories are not deterministic does not rob them of explanatory power. Nowadays most philosophers would agree that they have the capacity to explain statistical regularities; the question is whether they can be used to explain individual occurrences. If Principle I—which is *not* needed to block vacuous explanations—is relinquished we can give an affirmative answer.³⁷