not formulate, will need to be added to the D–N model in addition to his condition. This type of proposal will be examined in section 6 when Brody's causal model is discussed.

16. Kim, 'Causation, Nomic Subsumption, and the Concept of an Event', *Journal of Philosophy* 70 (1973), 217–236. This is a modification of Kim, whose formulations need a slight repair.

17. Wesley C. Salmon, *Statistical Explanation and Statistical Relevance* (Pittsburgh, 1971).

18. Salmon in his book does not explicitly say this, but this seems to be the most reasonable interpretation of his position. See e.g., his remarks on pp. 79–80 in which he is distinguishing homogeneity from epistemic homogeneity, and in which he compares his model with Hempel's. Salmon might, of course, say that he is supplying conditions only for the concept of a well-confirmed or justified explanation. But I am construing his model in a stronger sense, and in private conversation he assures me that this is the correct interpretation.

19. It is possible to construe Salmon's model as requiring the satisfaction of the homogeneity condition to be stated in the explanans itself. If so the model would purport to satisfy the *a priori* requirement. However, in what follows I shall continue to assume that only sentences of the type in (3) comprise the explanans, and thus that the model is an empirical one. On either construction the same difficulty will emerge.

20. A variety of such examples, as well as other trenchant criticism, can be found in John B. Meixner's doctoral dissertation, *Salmon's Statistical Model of Explanation* (The Johns Hopkins University, Baltimore, 1976).

21. B. A. Brody, 'Towards an Aristotelian Theory of Scientific Explanation', *Philosophy of Science* 39 (1972), pp. 20–31. There is a corresponding model for the explanation of laws, which I won't discuss.

22. See *ibid.*, p. 27.

23. If the explanans explicitly says that the property is essential, then the model purports to satisfy the *a priori* condition. But I want here to consider a model that explicitly violates this condition. (In either case the model turns out to be unsatisfactory.)

24. Brody, *op. cit.*, p. 23.

25. In his most recent writings, Salmon too proposes adding a causal condition to his S–R model. Unlike Brody and Woodward, he attempts to define the concept of causation he utilizes, although, by his own admission, the definition is not entirely adequate. See his 'Why ask "Why?"', *Proceedings and Addresses of the American Philosophical Association* 51 (1977), 683–705.

26. See C. J. Ducasse, 'Explanation, Mechanism, and Teleology', *Journal of Philosophy* XXII (1925), pp. 150–55; Donald Davidson, 'Actions, Reasons, and Causes', *Journal of Philosophy* LX (1963), pp. 685–700; Alvin I. Goldman, *A Theory of Human Action* Englewood Cliffs, N.J.), p. 78.

27. A cogent attack on Brody's model can be found in Timothy McCarthy, 'On an Aristotelian Model of Scientific Explanation', *Philosophy of Science* 44 (1977), 159–166. See Davidson, 'Psychology as Philosophy', in J. Glover (ed.), *Philosophy of Mind* (Oxford, 1976), pp. 103–104, for criticism of the causal-motivational model (which Davidson himself once supported).

28. Indeed, some formulations of the causal-motivational model (e.g., Ducasse's and Goldman's) seem to allow as an explicit part of the explanans a singular explanation-sentence that entails the explanandum. In such cases the NES requirement is violated in letter as well as in spirit.

29. Material in this paper is from a book I am writing on the nature of explanation. I am indebted to the National Science Foundation for support of research.

Bas C. van Fraassen
THE PRAGMATICS OF EXPLANATION

There are two problems about scientific explanation. The first is to describe it: when is something explained? The second is to show why (or in what sense) explanation is a virtue. Presumably we have no explanation unless we have a good theory; one which is independently worthy of acceptance. But what virtue is there in explanation over and above this? I believe that philosophical concern with the first problem has been led thoroughly astray by mistaken views on the second.

I. FALSE IDEALS

To begin I wish to dispute three ideas about explanation that seem to have a subliminal influence on the discussion. The first is that explanation is a relation simply between a theory or hypothesis and the phenomena or facts, just like truth for example. The second is that explanatory power cannot be logically separated from certain other virtues of a theory, notably truth or acceptability. And the third is that explanation is the overriding virtue, the end of scientific inquiry.

When is something explained? As a foil to the above three ideas, let me propose the simple answer: *when we have a theory which explains*. Note first that "have" is not "have on the books"; I cannot claim to have such a theory without implying that this theory is acceptable all told. Note also that both "have" and "explains" are tensed; and that I have allowed that we can have a theory

which does not explain, or "have on the books" an unacceptable one that does. Newton's theory explained the tides but not the advance in the perihelion of mercury; we used to have an acceptable theory, provided by Newton, which bore (or bears timelessly?) the explanation relationship to some facts but not to all. My answer also implies that we can intelligibly say that the theory explains, and not merely that people can explain by means of the theory. But this consequence is not very restrictive, because the former could be an ellipsis for the latter.

There are questions of usage here. I am happy to report that the history of science allows systematic use of both idioms. In Huygens and Young the typical phrasing seemed to be that phenomenon may be explained *by means of* principles, laws and hypotheses, or *according to* a view.[1] On the other hand, Fresnel writes to Arago in 1815 "Tous ces phénomènes . . . sont réunis et expliqués par la même théorie des vibrations," and Lavoisier says that the oxygen hypothesis he proposes *explains* the phenomena of combustion.[2] Darwin also speaks in the latter idiom: "In scientific investigations it is permitted to invent any hypothesis, and if it explains various large and independent classes of facts it rises to the rank of a well-grounded theory"; though elsewhere he says that the facts of geographical distribution are *explicable on* the theory of migration.[3]

My answer did separate acceptance of the theory from its explanatory power. Of course the second can be a reason for the first; but *that* requires their separation. Various philosophers have held that explanation logically requires true (or acceptable) theo-

ries as premises. Otherwise, they hold, we can at most mistakenly believe that we have an explanation.

This is also a question of usage, and again usage is quite clear. Lavoisier said of the phlogiston hypothesis that it is too vague and consequently "s'adapte a toutes les explications dans lesquelles on veut le faire entrer."[4] Darwin explicitly allows explanations by false theories when he says "It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified."[5] More recently, Gilbert Harman has argued similarly: that a theory explains certain phenomena is part of the evidence that leads us to accept it. But that means that the explanation-relation is visible beforehand. Finally, we criticize theories selectively: a discussion of celestial mechanics around the turn of the century would surely contain the assertion that Newton's theory does explain many planetary phenomena, though not the advance in the perihelion of Mercury.

There is a third false ideal, which I consider worst: that explanation is the *summum bonum* and exact aim of science. A virtue could be overriding in one of two ways. The first is that it is a minimal criterion of acceptability. Such is consistency with the facts in the domain of application (though not necessarily with all data, if these are dubitable!). Explanation is not like that, or else a theory would not be acceptable at all unless it explained all facts in its domain. The second way in which a virtue may be overriding is that of being required when it can be had. This would mean that if two theories pass other tests (empirical adequacy, simplicity) equally well, then the one which explains more must be accepted. As I have argued elsewhere,[6] and as we shall see in connection with Salmon's views below, a precise formulation of this demand requires hidden variables for indeterministic theories. But of course, hidden variables are rejected in scientific practice as so much "metaphysical

baggage" when they make no difference in empirical predictions.

II. A BIASED HISTORY

I will outline the attempts to characterize explanation of the past three decades, with no pretense of objectivity. On the contrary, the selection is meant to illustrate the diagnosis, and point to the solution, of the next section.

1. Hempel

In 1966, Hempel summarized his views by listing two main criteria for explanation. The first is the criterion of *explanatory relevance:* "the explanatory information adduced affords good grounds for believing that the phenomenon to be explained did, or does, indeed occur."[7] That information has two components, one supplied by the scientific theory, the other consisting of auxiliary factual information. The relationship of providing good grounds is explicated as (a) implying (D–N case), or (b) conferring a high probability (I–S case), which is not lowered by the addition of other (available) evidence.

As Hempel points out, this criterion is not a sufficient condition for explanation: the red shift gives us good grounds for believing that distant galaxies are receding from us, but does not explain why they do. The classic case is the *barometer example:* the storm will come exactly if the barometers fall, which they do exactly if the atmospheric conditions are of the correct sort; yet only the last factor explains. Nor is the criterion a necessary condition; for this the classic case is the *paresis example.* We explain why the mayor, alone among the townsfolk, contracted paresis by his history of latent, contracted syphilis; yet such histories are followed by paresis in only a small percentage of cases.

The second criterion is the requirement

of *testability;* but since all serious candidates for the role of scientific theory meet this, it cannot help to remove the noted defects.

2. Beckner, Putnam, and Salmon

The criterion of explanatory relevance was revised in one direction, informally by Beckner and Putnam and precisely by Salmon. Morton Beckner, in his discussion of evolution theory, pointed out that this often explains a phenomenon only by showing how it could have happened, given certain possible conditions.[8] Evolutionists do this by constructing models of processes which utilize only genetic and natural selection mechanisms, in which the outcome agrees with the actual phenomenon. Parallel conclusions were drawn by Hilary Putnam about the way in which celestial phenomena are explained by Newton's theory of gravity: celestial motions could indeed be as they are, given a certain possible (though not, known) distribution of masses in the universe.[9]

We may take the paresis example to be explained similarly. Mere consistency with the theory is of course much too weak, since that is implied by logical irrelevance. Hence Wesley Salmon made this precise as follows: to explain is to exhibit (the) statistically relevant factors.[10] (I shall leave till later the qualifications about "screening off.") Since this sort of explication discards the talk about modelling and mechanisms of Beckner and Putnam, it may not capture enough. And indeed, I am not satisfied with Salmon's arguments that his criterion provides a sufficient condition. He gives the example of an equal mixture of Uranium 238 atoms and Polonium 214 atoms, which makes the Geiger counter click in interval (t, $t + m$). This means that one of the atoms disintegrated. Why did it? The correct answer will be: because it was a Uranium 238 atom, if that is so—although the probability of its disintegration is much higher relative to the previous knowledge that the

atom belonged to the described mixture.[11] The problem with this argument is that, on Salmon's criterion, we can explain not only why there was a disintegration, but also why *that* atom disintegrated *just then.* And surely that is exactly one of those facts which atomic physics leaves unexplained.

But there is a more serious general criticism. Whatever the phenomenon is, we can amass the statistically relevant factors, as long as the theory does not rule out the phenomenon altogether. "What more could one ask of an explanation?" Salmon inquires.[12] But in that case, as soon as we have an empirically adequate theory, we have an explanation of every fact in its domain. We may claim an explanation as soon as we have shown that the phenomenon can be embedded in some model allowed by the theory— that is, does not throw doubt on the theory's empirical adequacy.[13] But surely that is too sanguine?

3. Global Properties

Explanatory power cannot be identified with empirical adequacy; but it may still reside in the performance of the theory as a whole. This view is accompanied by the conviction that science does not explain individual facts but general regularities and was developed in different ways by Michael Friedman and James Greeno. Friedman says explicitly that in his view, "the kind of understanding provided by science is global rather than local" and consists in the simplification and unification imposed on our world picture.[14] That S_1 explains S_2 is a conjunction of two facts: S_1 implies S_2 relative to our background knowledge (and/or belief) K, *and* S_1 unifies and simplifies the set of its consequences relative to K. Friedman will no doubt wish to weaken the first condition in view of Salmon's work.

The precise explication Friedman gives of the second condition does not work, and is not likely to have a near variant that does.[15] But here we may look at Greeno's

proposal.[16] His abstract and closing statement subscribe to the same general view as Friedman. But he takes as his model of a theory one which specifies a single probability space Q as the correct one, plus two partitions (or random variables) of which one is designated *explanandum* and the other one is designated *explanans*. An example: sociology cannot explain why Albert, who lives in San Francisco and whose father has a high income, steals a car. Nor is it meant to. But it does explain delinquency in terms of such other factors as residence and parental income. The degree of explanatory power is measured by an ingeniously devised quantity which measures the information I the theory provides of the explanandum variable M on the basis of explanans S. This measure takes its maximum value if all conditional probabilities $P(M_i/S_j)$ are zero or one (D–N case), and its minimum value zero if S and M are statistically independent.

Unfortunately, this way of measuring the unification imposed on our data abandons Friedman's insight that scientific understanding cannot be identified as a function of grounds for rational expectation. For if we let S and M describe the behavior of the barometer and coming storms, with $P(\text{barometer falls}) = P(\text{storm comes}) = 0.2$, $P(\text{storm comes/barometer falls}) = 1$, and $P(\text{storm comes/barometer does not fall}) = 0$, then the quantity I takes its maximum value. Indeed, it does so whether we designate M or S as explanans.

It would seem that such asymmetries as exhibited by the red shift and barometer examples must necessarily remain recalcitrant for any attempt to strengthen Hempel's or Salmon's criteria by global restraints on theories alone.

4. The Major Difficulties

There are two main difficulties, illustrated by the old paresis and barometer examples, which none of the examined positions can handle. The first is that there are cases, clearly in a theory's domain, where

the request for explanation is nevertheless rejected. We can explain why John, rather than his brothers contracted paresis, for he had syphilis; but not why he, among all those syphilitics, got paresis. Medical science is incomplete, and hopes to find the answer some day. But the example of the uranium atom disintegrating just then rather than later, is formally similar and we believe the theory to be complete. We also reject such questions as the Aristotelians asked the Galileans: why does a body free of impressed forces retain its velocity? The importance of this sort of case, and its pervasive character, has been repeatedly discussed by Adolf Grünbaum.

The second difficulty is the asymmetry revealed by the barometer: even if the theory implies that one condition obtains when and only when another does, it may be that and only when another does, it may be that it explains the one in terms of the other and not vice versa. An example which combines both the first and second difficulty is this: according to atomic physics, each chemical element has a characteristic atomic structure and a characteristic spectrum (of light emitted upon excitation). Yet the spectrum is explained by the atomic structure, and the question why a substance has that structure does not arise at all (except in the trivial sense that the questioner may need to have the terms explained to him).

5. Causality

Why are there no longer any Tasmanian natives? Well, they were a nuisance, so the white settlers just kept shooting them till there were none left. The request was not for population statistics, but for the story; though in some truncated way, the statistics "tell" the story.

In a later paper Salmon gives a primary place to causal mechanisms in explanation.[17] Events are bound into causal chains by two relations: spatio-temporal continuity and statistical relevance. Explanation requires the exhibition of such chains. Salmon's point of departure is Reichen-

bach's *principle of the common cause:* every relation of statistical relevance ought to be explained by one of causal relevance. This means that a correlation of simultaneous values must be explained by a prior common cause. Salmon gives two statistical conditions that must be met by a common cause C of events A and B:

(a) $P(A \text{ \& } B/C) = P(A/C)P(B/C)$
(b) $P(A/B \text{ \& } C) = P(A/C)$ "C *screens off B from A.*"

If $P(B/C) \neq 0$ these are equivalent, and symmetric in A and B.

Suppose that explanation is typically the demand for a common cause. Then we still have the problem: when does this arise? Atmospheric conditions explain the correlation between barometer and storm, say; but are still prior causes required to explain the correlation between atmospheric conditions and falling barometers?

In the quantum domain, Salmon says, causality is violated because "causal influence is not transmitted with spatio-temporal continuity." But the situation is worse. To assume Reichenbach's principle to be satisfiable, continuity aside, is to rule out all genuinely indeterministic theories. As example, let a theory say that C is invariably followed by one of the incompatible events A, B, or D, each with probability $\frac{1}{3}$. Let us suppose the theory complete, and its probabilities irreducible, with C the complete specification of state. Then we will find a correlation for which only C could be the common cause, but it is not. Assuming that A, B, D are always preceded by C and that they have low but equal prior probabilities, there is a statistical correlation between $\phi = (A \text{ or } D)$ and $\psi = (B \text{ or } D)$, for $P(\phi/\psi) = P(\psi/\phi) = \frac{1}{2} \neq P(\phi)$. But C, the only available candidate, does not screen off ϕ from ψ: $P(\phi/C \text{ \& } \psi) = P(\phi/\psi) = \frac{1}{2} \neq P(\phi/C)$ which is $\frac{2}{3}$. Although this may sound complicated, the construction is so general that almost any irreducibly probabilistic situation will give a similar example. Thus Reichenbach's *princi-*

ple of the common cause is in fact a demand for hidden variables.

Yet we retain the feeling that Salmon has given an essential clue to the asymmetries of explanation. For surely the crucial point about the barometer is that the atmospheric conditions screen off the barometer fall from the storm? The general point that the asymmetries are totally bound up with causality was argued in a provocative article by B. A. Brody.[18] Aristotle certainly discussed examples of asymmetries: the planets do not twinkle because they are near, yet they are near if and only if they do not twinkle (*Posterior Analytics*, I, 13). Not all explanations are causal, says Brody, but the others use a second Aristotelian notion, that of essence. The spectrum angle is a clear case: sodium has that spectrum because it has this atomic structure, which is its essence.

Brody's account has the further advantage that he can say when questions do not arise: other properties are explained in terms of essence, but the request for an explanation of the essence does not arise. However, I do not see how he would distinguish between the questions why the uranium atom disintegrated and why it disintegrated just then. In addition there is the problem that modern science is not formulated in terms of causes and essences, and it seems doubtful that these concepts can be redefined in terms which do occur there.

6. Why-Questions

A why-question is a request for explanation. Sylvain Bromberger called P the *presupposition* of the question *Why-P?* and restated the problem of explanation as that of giving the conditions under which proposition Q is a correct answer to a why-question with presupposition P.[19] However, Bengt Hannson has pointed out that "Why was it John who ate the apple?" and "Why was it the apple which John ate?" are different why-questions, although the com-

prised proposition is the same.[20] The difference can be indicated by such phrasing, or by emphasis ("Why did *John* . . . ?") or by an auxiliary clause ("Why did John rather than . . . ?"). Hannson says that an explanation is requested, not of a proposition or fact, but of an *aspect* of a proposition.

As is at least suggested by Hannson, we can cover all these cases by saying that we wish an explanation of why *P* is true in contrast to other members of a set *X* or propositions. This explains the tension in our reaction to the paresis-example. The question why the mayor, in contrast to other townfolk generally, contracted paresis *has* a true correct answer: because of his latent syphilis. But the question why he did in contrast to the other syphilitics in his country club, has no true correct answer. Intuitively we may say: *Q* is a correct answer to *Why P in contrast to X?* only if *Q* gives reasons to expect that *P*, in contrast to the other members of *X*. Hannson's proposal for a precise criterion is: the probability of *P* given *Q* is higher than the average of the probabilities of *R* given *Q*, for members *R* of *X*.

Hannson points out that the set *X* of alternatives is often left tacit; the two questions about paresis might well be expressed by the same sentence in different contexts. The important point is that explanations are not requested of propositions, and consequently a distinction can be drawn between answered and rejected requests in a clear way. However, Hannson makes *Q* a correct answer to *Why P in contrast to X?* when *Q* is statistically irrelevant, when *P* is already more likely than the rest; or when *Q* implies *P* but not the others. I do not see how he can handle the barometer (or red shift, or spectrum) asymmetries. On his precise criterion, that the barometer fell is a correct answer to why it will storm as opposed to be calm. The difficulty is very deep: if *P* and *R* are necessarily equivalent, according to our accepted theories, how can *Why P in contrast to X?* be distinguished from *Why R in contrast to X?*

III. THE SOLUTION

1. Prejudices

Two convictions have prejudiced the discussion of explanation, one methodological and one substantive.

The first is that a philosophical account must aim to produce necessary and sufficient conditions for theory *T* explaining phenomenon *E*. A similar prejudice plagued the discussion of counter-factuals for twenty years, requiring the exact conditions under which, if *A* were the case, *B* would be. Stalnaker's liberating insight was that these conditions are largely determined by context and speaker's interest. This brings the central question to light: what *form* can these conditions take?

The second conviction is that explanatory power is a virtue of theories by themselves, or of their relation to the world, like simplicity, predictive strength, truth, empirical adequacy. There is again an analogy with counterfactuals: it used to be thought that science contains, or directly implies, counterfactuals. In all but limiting cases, however, the proposition expressed is highly context-dependent, and the implication is there at most relative to the determining contextual factors, such as speakers' interest.

2. Diagnosis

The earlier accounts lead us to the format: *C* explains *E* relative to theory *T* exactly if (a) *T* has certain global virtues, and (b) *T* implies a certain proposition $\phi(C,E)$ expressible in the language of logic and probability theory. Different accounts directed themselves to the specification of what should go into (a) and (b). We may add, following Beckner and Putnam, that *T* explains *E* exactly if there is a proposition *C* consistent with *T* (and presumably, background beliefs) such that *C* explains *E* relative to *T*.

The significant modifications were proposed by Hannson and Brody. The former pointed out that the explanadum *E* cannot be reified as a proposition: we request the explanation of something *F* in contrast to its alternatives *X* (the latter generally tacitly specified by context). This modification is absolutely necessary to handle some of our puzzles. It requires that, in (b) above, we replace "$\phi(C,E)$" by the formula form "$\psi(C,F,X)$." But the problem of asymmetries remains recalcitrant, because if *T* implies the necessary equivalence of *F* and *F'* (say, atomic structure and characteristic spectrum), then *T* will also imply $\psi(C,F',X)$ if and only if it implies $\psi(C,F,X)$.

The only account we have seen which grapples at all successfully with this, is Brody's. For Brody points out that even properties which we believe to be constantly conjoined in all possible circumstances, can be divided into essences and accidents, or related as cause and effect. In this sense, the asymmetries were no problem for Aristotle.

3. The Logical Problem

We have now seen exactly what logical problem is posed by the asymmetries. To put it in current terms: how can we distinguish propositions which are true in exactly the same possible worlds?

There are several known approaches that use impossible worlds. David Lewis, in his discussion of causality, suggests that we should look not only to the worlds theory *T* allows as possible, but also to those it rules out as impossible, and speaks of counterfactuals which are counterlegal. Relevant logic and entailment draw distinctions between logically equivalent sentences and their semantics devised by Routley and Meyer use both inconsistent and incomplete worlds. I believe such approaches to be totally inappropriate for the problem of explanation, for when we look at actual explanations of phenomena by theories, we do not see any

detours through circumstances or events ruled out as impossible by the theory.

A further approach, developed by Rolf Schock, Romane Clark, and myself distinguishes sentences by the facts that make them true. The idea is simple. That it rains, that it does not rain, that it snows, and that it does not snow, are four distinct facts. The disjunction that it rains or does not rain is made true equally by the first and second, and not by the third or fourth, which distinguishes it from the logically equivalent disjunction that it snows or does not snow.[21] The distinction remains even if there is also a fact of its raining or not raining, distinct or identical with that of its snowing or not snowing.

This approach can work for the asymmetries of explanation. Such asymmetries are possible because, for example, the distinct facts that light is emitted with wavelengths λ, μ, . . . conjointly make up the characteristic spectrum, while quite different facts conjoin to make up the atomic structure. So we have shown how such asymmetries *can* arise, in the way that Stalnaker showed how failures of transitivity in counterfactuals *can* arise. But while we have the distinct facts to classify asymmetrically, we still have the non-logical problem: whence comes the classification? The only suggestion so far is that it comes from Aristotle's concepts of cause and essence; but if so, modern science will not supply it.

4. The Aristotelian Sieve

I believe that we should return to Aristotle more thoroughly, and in two ways. To begin, I will state without argument how I understand Aristotle's theory of science. Scientific activity is divided into two parts, *demonstration* and *explanation*, the former treated mainly by the *Posterior Analytics* and the latter mainly by Book II of the *Physics*. Illustrations in the former are mainly examples of explanations in which the results of demonstration are *applied*; this is why the examples

contain premises and conclusions which are not necessary and universal principles, although demonstration is only to and from such principles. Thus the division corresponds to our pure versus applied science. There is no reason to think that principles and demonstrations have such words as "cause" and "essence" in them, although looking at pure science from outside, Aristotle could say that its principles state causes and essences. In applications, the principles may be filtered through a conceptual sieve originating outside science.

The doctrine of the four "causes" (*aitiai*) allows for the systematic ambiguity or context-dependence of why-questions.[22] Aristotle's example (Physics II, 3; 195a) is of a lantern. In a modern example, the question why the porch light is on may be answered "because I flipped the switch" or "because we are expecting company," and the context determines which is appropriate. Probabilistic relations cannot distinguish these. Which factors are explanatory is decided not by features of the scientific theory but by concerns brought from outside. This is true even if we ask specifically for an "efficient cause," for how far back in the chain should we look, and which factors are merely auxiliary contributors?

Aristotle would not have agreed that essence is context-dependent. The essence is what the thing *is*, hence, its sum of classificatory properties. Realism has always asserted that ontological distinctions determine the "natural" classification. But which property is counted as explanatory and which as explained seems to me clearly context dependent. For consider Bromberger's flagpole example: the shadow is so long because the pole has this height, and not conversely. At first sight, no contextual factor could reverse this asymmetry, because the pole's height is a property it has in and by itself, and its shadow is a very accidental feature. The general principle linking the two is that its shadow is a function $f(x,t)$ of its height x and the time t (the latter determining the sun's elevation). But imagine the

pole is the pointer on a giant sundial. Then the values of f have desired properties for each time t, and we appeal to these to explain why it is (had to be) such a tall pole.

We may again draw a parallel to counterfactuals. Professor Geach drew my attention to the following spurious argument: If John asked his father for money, then they would not have quarreled (because John is too proud to ask after a quarrel). Also if John asked and they hadn't quarreled, he would receive. By the usual logic of counterfactuals, it follows that if John asked his father for money, he would receive. But we know that he would not, because they have in fact quarreled. The fallacy is of equivocation, because "what was kept constant" changed in the middle of the monologue. (Or if you like, the aspects by which worlds are graded as more or less similar to this one.) Because science cannot dictate what speakers decide to "keep constant" it contains no counterfactuals. By exact parallel, *science contains no explanations.*

5. The Logic of Why-Questions

What remains of the problem of explanation is to study its logic, which is the logic of why-questions. This can be put to some extent, but not totally, in the general form developed by Harrah and Belnap and others.[23]

A question admits of three classes of response, *direct answers, corrections,* and *comments.* A *presupposition,* it has been held, is any proposition implied by all direct answers, or equivalently, denied by a correction. I believe we must add that the question "Why P, in contrast to X?" also presupposes that (a) P is a member of X, (b) P is true and the majority of X are not. This opens the door to the possibility that a question may not be uniquely determined by its set of direct answers. The question itself should decompose into factors which determine that set: the *topic P,* the *alternative X,* and a *request specification* (of which the doctrine of

the four "causes" is perhaps the first description).

We have seen that the propositions involved in question and answer must be individuated by something more than the set of possible worlds. I propose that we use the facts that make them true (see footnote 21). The context will determine an asymmetric relation among these facts, of *explanatory relevance;* it will also determine the theory or beliefs which determine which worlds are *possible,* and what is *probable* relative to what.

We must now determine what direct answers are and how they are evaluated. They must be made true by facts (and only by facts forcing such) which are explanatorily relevant to those which make the topic true. Moreover, these facts must be statistically relevant, telling for the topic in contrast to the alternatives generally; this part I believe to be explicable by probabilities, combining Salmon's and Hannson's account. How strongly the answers count for the topic should be part of their evaluation as better or worse answers.

The main difference from such simple questions as "Which cat is on the mat?" lies in the relation of a why-question to its presuppositions. A why-question may fail to arise because it is ill-posed (P is false, or most of X is true), or because only question-begging answers tell probabilistically for P in contrast to X generally, or because none of the factors that do tell for P are explanatorily relevant in the question-context. Scientific theory enters mainly in the evaluation of possibilities and probabilities, which is only part of the process, and which it has in common with other applications such as prediction and control.

IV. SIMPLE PLEASURES

There are no explanations in science. How did philosophers come to mislocate explanation among semantic rather than pragmatic relations? This was certainly in part because

the positivists tended to identify the pragmatic with subjective psychological features. They looked for measures by which to evaluate theories. Truth and empirical adequacy are such, but they are weak, being preserved when a theory is watered down. Some measure of "goodness of fit" was also needed, which did not reduce to a purely internal criterion such as simplicity, but concerned the theory's relation to the world. The studies of explanation have gone some way toward giving us such a measure, but it was a mistake to call this explanatory power. The fact that seemed to confirm this error was that we do not say that we *have* an explanation unless we have a theory which is acceptable, and victorious in its competition with alternatives, whereby we can explain. Theories are applied in explanation, but the peculiar and puzzling features of explanation are supplied by other factors involved. I shall now redescribe several familiar subjects from this point of view.

When a scientist campaigns on behalf of an advocated theory, he will point out how our situation will change if we accept it. Hitherto unsuspected factors become relevant, known relations are revealed to be strands of an intricate web, some terribly puzzling questions are laid to rest as not arising at all. We shall be in a much better position to explain. But equally, we shall be in a much better position to predict and control. The features of the theory that will make this possible are its empirical adequacy and logical strength, not special "explanatory power" and "control power." On the other hand, it is also a mistake to say explanatory power is nothing but those other features, for then we are defeated by asymmetries having no "objective" basis in science.

Why are *new* predictions so much more to the credit of a theory than agreement with the old? Because they tend to bring to light new phenomena which the older theories cannot explain. But of course, in doing so, they throw doubt on the empirical adequacy of the older theory: they show that a pre-

condition for explanation is not met. As Boltzmann said of the radiometer, "the theories based on older hydrodynamic experience can never describe" these phenomena.[24] The failure in explanation is a by-product.

Scientific inference is inference to the best explanation. That does not rule at all for the supremacy of explanation among the virtues of theories. For we evaluate how good an explanation is given by how good a theory is used to give it, how close it fits to the empirical facts, how internally simple and coherent the explanation. There is a further evaluation in terms of a prior judgment of which kinds of factors are explanatorily relevant. If this further evaluation took precedence, overriding other considerations, explanation would be the peculiar virtue sought above all. But this is not so: instead, science schools our imagination so as to revise just those prior judgments of what satisfies and eliminates wonder.

Explanatory power is something we value and desire. But we are as ready, for the sake of scientific progress, to dismiss questions as not really arising at all. Explanation is indeed a virtue; but still, less a virtue than an anthropocentric pleasure.[25]

NOTES

1. I owe these and following references to my student Mr. Paul Thagard. For instance see C. Huygens, *Treatise on Light*, tr. by S. P. Thompson (New York, 1962), pp. 19, 20, 22, 63; Thomas Young, *Miscellaneous Works*, ed. by George Peacock (London, 1855), Vol. I, pp. 168, 170.

2. Augustin Fresnel, *Oeuvres Complètes* (Paris, 1866), Vol. I, p. 36 (see also pp. 254, 355); Antoine Lavoisier, *Oeuvres* (Paris, 1862), Vol. II, p. 233.

3. Charles Darwin, *The Variation of Animals and Plants* (London, 1868), Vol. 1, p. 9; *On the Origin of the Species* (Facs. of first edition, Cambridge, Mass., 1964), p. 408.

4. Antoine Lavoisier, *op. cit.*, p. 640.

5. *Origin* (sixth ed., New York, 1962), p. 476.

6. "Wilfrid Sellars on Scientific Realism," *Dialogue*, Vol. 14 (1975), pp. 606–616.

7. C. G. Hempel, *Philosophy of Natural Science* (Englewood Cliffs, New Jersey, 1966), p. 48.

8. *The Biological Way of Thought* (Berkeley, 1968), p. 176; this was first published in 1959.

9. In a paper of which a summary is found in Frederick Suppe (ed.), *The Structure of Scientific Theories* (Urbana, Ill., 1974).

10. "Statistical Explanation," pp. 173–231 in R. G. Colodny (ed.), *The Nature and Function of Scientific Theories* (Pittsburgh, 1970); reprinted also in Salmon's book cited below.

11. *Ibid.*, pp. 207–209. Nancy Cartwright has further, unpublished, counter-examples to the necessity and sufficiency of Salmon's criterion.

12. *Ibid.*, p. 222.

13. These concepts are discussed in my "To Save the Phenomena," *The Journal of Philosophy*, Vol. 73 (1976), forthcoming.

14. "Explanation and Scientific Understanding," *The Journal of Philosophy*, Vol. 71 (1974), pp. 5–19.

15. See Philip Kitcher, "Explanation, Conjunction, and Unification," *The Journal of Philosophy*, Vol. 73 (1976), pp. 207–212.

16. "Explanation and Information," pp. 89–103 in Wesley Salmon (ed.), *Statistical Explanation and Statistical Relevance* (Pittsburgh, 1971). This paper was originally published with a different title in *Philosophy of Science*, Vol. 37 (1970), pp. 279–293.

17. "Theoretical Explanation," pp. 118–145 in Stephan Körner (ed.), *Explanation* (Oxford, 1975).

18. "Towards an Aristotelian Theory of Scientific Explanation," *Philosophy of Science*, Vol. 39 (1972), pp. 20–31.

19. "Why-Questions," pp. 86–108 in R. G. Colodny (ed.), *Mind and Cosmos* (Pittsburgh, 1966).

20. "Explanations—Of What?" (mimeographed: Stanford University, 1974).

21. Cf. my "Facts and Tautological Entailments," *The Journal of Philosophy*, Vol. 66 (1969), pp. 477–487 and in A. R. Anderson, *et al*, (ed.), *Entailment* (Princeton, 1975); and "Extension, Intension, and Comprehension" in Milton Munitz (ed.), *Logic and Ontology* (New York, 1973).

22. Cf. Julius Moravcik, "Aristotle on Adequate Explanations," *Synthese*, Vol. 28 (1974), pp. 3–18.

23. Cf. N. D. Belnap, Jr., "Questions: Their Presuppositions, and How They Can Fail to Arise," *The Logical Way of Doing Things*, ed. by Karel Lambert (New Haven, 1969), pp. 23–39.

24. Ludwig Boltzmann, *Lectures on Gas Theory*, tr. by S. G. Brush (Berkeley, 1964), p. 25.

25. The author wishes to acknowledge helpful discussions and correspondence with Professors N. Cartwright, B. Hannson, K. Lambert, and W. Salmon, and the financial support of the Canada Council.

The Nature of Causality

Michael Scriven
CAUSATION AS EXPLANATION

0. INTRODUCTION

Some of the most widely used concepts in the logical framework of science—explanation, causation, and evaluation would seem to be good candidates for that honor—enjoy a curious love-hate relationship with philosophers of science. Though many philosophers have made serious and diverse efforts to analyze them, there have been others, equally eminent, who have completely denied their importance or their legitimacy. It has been argued that the task of science is to describe and not to explain; that reference to causation is a sign of an immature science and not to be found in the more advanced parts of the physical sciences today; and that value-free science should be regarded as the appropriate ideal.

I believe that in each case, the arguments for the rejection are fallacious and not even very subtly fallacious but I would agree that there is one great attraction about the nihilistic positions, which is that they avoid the exceedingly difficult task of providing a reasonable analysis of the concepts in question. It is even possible that the extraordi-

nary recalcitrance of these concepts has driven or at least assisted philosophers to conclude that they are not respectable.

Social respectability is a matter of the directness with which one can be linked to indubitably respectable antecedents, and philosophical respectability is a matter of the directness of the linkage to philosophically respectable components. Given the rather limited group of respectable basic concepts with which positivist and neopositivist philosophers of science began, there arose a rather snobbish view of a good many concepts in the scientist's working vocabulary. From my point of view, the time for this aprioristic approach has long gone, and I have spent a great many years arguing for a more plebeian approach to the philosophy of science, an approach which analyzes scientific language without reference to any social register from previous epistemological excursions. Sometimes this leads to its own, hopefully more natural, reductions, and I believe this is the case with the three concepts I have mentioned here.

A somewhat general way to put the point would be to say that these concepts are all pragmatic rather than syntactic or semantic concepts; that their function can only be understood or explicated by reference to a