The Journal of Philosophy, Inc.

Explanation and Scientific Understanding
Author(s): Michael Friedman
Source: *The Journal of Philosophy*, Vol. 71, No. 1 (Jan. 17, 1974), pp. 5-19
Published by: The Journal of Philosophy, Inc.
Stable URL: https://www.jstor.org/stable/2024924
Accessed: 12-03-2026 21:40 UTC

THE JOURNAL OF PHILOSOPHY

VOLUME LXXI, NO. 1, JANUARY 17, 1974

EXPLANATION AND SCIENTIFIC UNDERSTANDING *

WHY does water turn to steam when heated? Why do the planets obey Kepler's laws? Why is light refracted by a prism? These are typical of the questions science tries to answer. Consider, for example, the answer to the first question: Water is made of tiny molecules in a state of constant motion. Between these molecules are intermolecular forces, which, at normal temperatures, are sufficient to hold them together. If the water is heated, however, the energy, and consequently the motion, of the molecules increases. If the water is heated sufficiently the molecules acquire enough energy to overcome the intermolecular forces—they fly apart and escape into the atmosphere. Thus, the water gives off steam. This account answers our question. Our little story seems to give us understanding of the process by which water turns to steam. The phenomenon is now more intelligible or comprehensible. How does this work? What is it about our little story, and scientific explanations generally, that gives us understanding of the world—what is it for a phenomenon to be scientifically understandable?

Two aspects of our example are of special interest. First, what is explained is a general regularity or pattern of behavior—a law, if you like—i.e., that water turns to steam when heated. Although most of the philosophical literature on explanation deals with the explanation of particular events, the type of explanation illustrated by the above account seems much more typical of the physical sciences. Explanations of particular events are comparatively rare—found only perhaps in geology and astronomy. Second, our little story explains one phenomenon, the changing of water into steam,

---

* An earlier version of this paper was read at the University of Massachusetts at Amherst. I am indebted to members of the philosophy department there, especially Fred Feldman, for helpful comments. I would also like to thank Hartry Field, Clark Glymour, and David Hills for valuable criticism and conversation.

by relating it to another phenomenon, the behavior of the molecules of which water is composed. This relation is commonly described as *reduction:* the explained phenomenon is said to be reduced to the explaining phenomenon; e.g., the behavior of water is reduced to the behavior of molecules. Thus, the central problem for the theory of scientific explanation comes down to this: what is the relation between phenomena in virtue of which one phenomenon can constitute an explanation of another, and what is it about this relation that gives understanding of the explained phenomenon?

When I ask that a theory of scientific explanation tell us what it is about the explanation relation that produces understanding, I do not suppose that 'scientific understanding' is a clear notion. Nor do I suppose that it is possible to say what scientific understanding is in advance of giving a theory of explanation. It is not reasonable to require that a theory of explanation proceed by first defining 'scientific understanding' and then showing how its reconstruction of the explanation relation produces scientific understanding. We can find out what scientific understanding consists in only by finding out what scientific explanation is and vice versa. On the other hand, although we have no clear independent notion of scientific understanding, I think we do have some general ideas about what features such a notion should have, and we can use these general ideas to judge the adequacy of philosophical theories of explanation. At any rate, this is the method I will follow. I will argue that traditional accounts of scientific explanation result in notions of scientific understanding that have objectionable or counterintuitive features. From my discussion of traditional theories I will extract some general properties that a concept of scientific understanding ought to have. Finally, I will propose an account of scientific explanation that possesses these desirable properties.

It seems to me that the philosophical literature on explanation falls into two rough groups. Some philosophers, like Hempel and Nagel, have relatively precise proposals as to the nature of the explanation relation, but relatively little to say about the connection between their proposals and scientific understanding, i.e., about what it is about the relation they propose that gives us understanding of the world. Other philosophers, like Toulmin, Scriven, and Dray, have a lot to say about understanding, but relatively vague ideas about just what relation it is that produces this understanding. To illustrate this situation I will discuss three attempts at explicating the explanation relation that have been prominent in the literature.

1. The best known philosophical account of explanation, the D-N model, is designed primarily as a theory about the explanation of particular events, but the view that the explanation relation is basically a deductive relation applies equally to our present concern. According to the D-N model, a description of one phenomenon can explain a description of a second phenomenon only if the first description entails the second. Of course, a deductive relation between two such descriptions is not sufficient for one to be an explanation of the other, as expounders of the D-N model readily admit.[1]

The entailment requirement puts a constraint on the explanation relation, but it does not by itself tell us what it is about the explanation relation that gives us understanding of the explained phenomenon, that makes the world more intelligible. In some of their writings defenders of the D-N model give the impression that they consider such a task to lie outside the province of the philosopher of science, because concepts like 'understanding' and 'intelligibility' are psychological or pragmatic. For example, Hempel writes: "such expressions as 'realm of understanding' and 'comprehensible' do not belong to the vocabulary of logic, for they refer to psychological or pragmatic aspects of explanation" (413). He goes on to characterize the pragmatic aspects of explanation as those which vary from individual to individual. Explanation in its pragmatic aspects is "a relative notion, something can be sgnificantly said to constitute an explanation in this sense only for this or that individual" (426). The philosopher of science, according to Hempel, aims at explicating the nonpragmatic aspects of explanation, the sense of 'explanation' on which *A* explains *B simpliciter* and not *for you or for me.*

I completely agree with Hempel's contention that the philosopher of science should be interested in an objective notion of explanation, a notion that doesn't vary capriciously from individual to individual. However, the considerations Hempel advances can serve as an argument against the attempt to connect explanation and understanding only by an equivocation on 'pragmatic'. In the sense in which such concepts as 'understanding', and 'comprehensible' are clearly pragmatic, 'pragmatic' means roughly the same as 'psychological', i.e., having to do with the thoughts, beliefs, attitudes,

---

[1] Compare C. G. Hempel and P. Oppenheim. "Studies in the Logic of Explanation," in Hempel's *Aspects of Scientific Explanation* (New York: Free Press, 1965), p. 273, note 33; parenthetical page references to Hempel are to this article. The difficulty is that the conjunction of two laws entails each of its conjuncts but does not necessarily explain them.

etc. of persons. However, 'pragmatic' can also mean subjective as opposed to objective. In this sense, a pragmatic notion varies from individual to individual, and is therefore a relative notion. But a concept can be pragmatic in the first sense without being pragmatic in the second. Take the concept of rational belief, for example—presumably, if it is rational to believe a given sentence on given evidence it is so for anyone, and not merely for this or that individual. Similarly, although the notion of understanding, like knowledge and belief but unlike truth, just is a psychological notion, I don't see why it can't be a perfectly objective one. I don't see why there can't be an objective or rational sense of 'scientific understanding', a sense on which what is scientifically comprehensible is constant for a relatively large class of people. Therefore, I don't see how the philosopher of science can afford to ignore such concepts as 'understanding' and 'intelligibility' when giving a theory of the explanation relation.

Despite his reluctance, Hempel as a matter of fact does try to connect his model of explanation with the notion of understanding. He writes: "the [D-N] argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred" (327). Here, showing that a phenomenon was to be expected comes to this: if one had known "the particular circumstances and laws in question" before the explained phenomenon occurred, one would have had rational grounds for expecting the explained phenomenon to occur. The phenomenon would not have taken one by surprise.

This attempt at connecting explanation and understanding is clearly best suited to the special case of the explanation of particular events; for only particular events occur at definite times, and can thus actually be expected before their occurrence. Nevertheless, the account seems to fail even here, since understanding and rational expectation are quite distinct notions. To have grounds for rationally expecting some phenomenon is not the same as to understand it. I think that this contention is conclusively established by the well-known examples of prediction via so-called "indicator laws"—the barometer and the storm, Koplick spots and measles, etc. In these examples, one is able to predict some phenomenon on the basis of laws and initial conditions, but one has not thereby enhanced one's understanding of why the phenomenon occurred. To the best of my knowledge, Hempel himself accepts these counterexamples, and, because of them, would concede today that the D-N model provides at best necessary conditions for the explanation of particular events.

When we come to the explanation of general regularities or patterns of behavior the situation is even worse. Since general regularities do not occur at definite times, there is no question of literally expecting them. In this context, showing a phenomenon to have been expected can only mean having rational grounds for believing that the phenomenon does occur. And it is clear that having grounds for believing that a phenomenon occurs, though it may be part of understanding that phenomenon, is not sufficient for such understanding. Scientific explanations may involve the provision of grounds for believing that the explained phenomena occur, but it is not in virtue of the provision of such grounds that they give us understanding.

I conclude that the D-N model has the following advantages: It provides a clear, precise, and simple condition—entailment—that the explanation relation must satisfy, which, as a necessary condition, is not obviously mistaken. Also, it makes explanation relatively objective—what counts as an explanation does not depend on the arbitrary tastes of the scientist or the age. However, D-N theorists have not succeeded in saying what it is about the explanation relation that provides understanding of the world.

2. A second view, which has been surprisingly popular, holds that scientific explanations give us understanding of the world by relating (or reducing) unfamiliar phenomena to familiar ones. This view is inspired by such examples as the kinetic theory of gases, which, it is thought, gain their explanatory power by comparing unfamiliar phenomena, such as the Boyle-Charles law, to familiar phenomena, such as the movements of tiny billiard balls. P. W. Bridgman states this view very clearly: "I believe that examination will show that the essence of explanation consists in reducing a situation to elements with which we are so familiar that we accept them as a matter of course, so our curiosity rests."[2] Among contemporary writers, William Dray seems to favor this view:

> Why does the theory of geometrical optics explain the length of particular shadows? . . . it is surely because a ray diagram goes along with it, allowing us to think of light as travelling along ray lines, some of the lines passing over the wall and others coming to a dead halt on its surface. The shadow length is explained when . . . we think of

---

[2] *The Logic of Modern Physics* (New York: Macmillan, 1968), p. 37.

> light as 'something travelling', i.e., when we apply to it a very familiar and perhaps anthropomorphic way of thinking. . . . Thus the role of theory in such explanations is really *parasitic* upon the fact that it suggests, with the aid of postulated, unobservable, entities, a 'hat-doffing' series of happenings which we are licensed to fill in . . . But if the travelling of observable entities along observable rails in a similar way would not explain a similar pattern of impact on encountering a wall, and if the jostling of a tightly packed crowd would not explain the straining and collapsing of the walls of a tent in which they were confined, then the corresponding scientific theories would not explain shadow lengths and the behavior of gases.[3]

The implication here is clearly that theories like the kinetic theory of gases are able to explain phenomena only to the extent that they relate them to more familiar processes and events.

This view, although it is initially attractive and does make an honest attempt to relate explanation and understanding, is rather obviously inadequate. As a matter of fact, many scientific explanations relate relatively familiar phenomena, such as the reflecting and refracting of light, to relatively unfamiliar phenomena, such as the behavior of electromagnetic waves. If the view under consideration were correct, most of the explanations offered by contemporary physics, which postulate phenomena stranger and less familiar than any that they explain, could not possibly explain. But, on reflection, it is not hard to see why this account fails so completely. For, being familiar, just like being expected, is not at all the same thing as being understood. We are all familiar with the behavior of household appliances like radios, televisions, and refrigerators; but how many of us understand why they behave the way they do?

Michael Scriven, although he explicitly rejects the "familiarity" account of explanation, appears to hold a view which is similar to it in important respects. He believes that in any given context each person possesses a "realm of understanding"—a set of phenomena which that person understands at a given time. The job of explanation is to relate phenomena that are not in this privileged set to phenomena that are in it:

> . . . the request for an explanation presupposes that *something* is understood, and a complete answer is one that relates the object of inquiry to the realm of understanding in some comprehensible and appropriate way. What this way is varies from subject matter to subject matter just as what makes something better than something else

---

[3] *Laws and Explanation in History* (New York: Oxford, 1964), pp. 79/80.

> varies from the field of automobiles to solutions of chess problems, but the *logical function* of explanation, as of evaluation, is the same in each field.[4]

Thus, whereas on the "familiarity" view of the explanation relation the phenomenon being explained must be related to a phenomenon that is familiar, on Scriven's account the phenomenon being explained must be related to a phenomenon that is already understood. On both views the phenomenon doing the explaining must have a special epistemological status. Not just any phenomenon will do. Both views conflict with the D-N account on this point. For, according to the D-N model, any phenomenon (regardless of its familiarity or epistemological status) that bears the appropriate deductive relation to the phenomenon being explained will do.

Scriven's view seems to me to be inadequate for the same reason as the "familiarity" view is. There just are many explanations in science which relate the phenomena being explained to phenomena that are not themselves understood in the relevant sense; i.e., we do not understand why these latter phenomena occur. This is true whenever a phenomenon is explained by reducing it to some "basic" or "fundamental" processes; e.g., an explanation in terms of the behavior of the fundamental particles of physics. In such cases the phenomenon doing the explaining is not itself understood; it is simply a brute fact. But its ability to explain *other* phenomena is not thereby impaired. Thus, I think that neither Scriven nor the "familiarity" theorists have given us good reason to suppose that it is a necessary feature of the explanation relation that the phenomenon doing the explaining must itself have some special epistemological status. It does not have to be a familiar or "hat doffing" phenomenon, nor do we have to understand why *it* occurs. It merely has to explain the phenomenon to which it is related.

3. A third approach to the explanation relation can be rather uncharitably labeled the "intellectual fashion" view. Like the "familiarity" theorists, holders of this view believe that the phenomenon doing the explaining must have a special epistemological status, but, unlike the "familiarity" theorists, they think that this status varies from scientist to scientist and from historical period to historical period. At any given time certain phenomena are regarded as

---

[4] "Explanations, Predictions, and Laws," in H. Feigl and G. Maxwell, eds., *Minnesota Studies in the Philosophy of Science*, vol. III (Minneapolis: Univ. of Minnesota Press, 1970), p. 202.

somehow self-explanatory or natural. Such phenomena need no explanation; they represent ideals of intelligibility. Explanation, within a particular historical tradition, consists in relating other phenomena to such ideals of intelligibility. Perhaps the clearest statement of this view is that of Stephen Toulmin, who calls such self-explanatory phenomena "ideals of natural order":

> . . . the scientist's prior expectations are governed by certain rational ideas or conceptions of the regular order of Nature. Things which happen according to these ideas he finds unmysterious; the cause or explanation of an event comes in question . . . through seemingly deviating from this regular way; its classification among the different sorts of phenomena (e.g., 'anomalous refraction') is decided by contrasting it with the regular, intelligible case; and, before the scientist can be satisfied, he must find some way of applying or extending or modifying his prior ideas about Nature so as to bring the deviant event into the fold.[5]

Thus, the meaning of 'scientific understanding' varies with historical tradition, since what counts as an ideal of intelligibility does. Consequently, the very same theory may count as explanatory for one tradition but may fail to explain for another.

In all fairness, it should be pointed out that most supporters of this account do not believe that the choice of such ideals of intelligibility is completely capricious, depending only on the whims and prejudices of particular scientists. On the contrary, most believe that there can be good reasons, usually having to do with predictive power, for choosing one ideal over another. Indeed, one writer, N. R. Hanson, practically identifies predictive power with intelligibility. He argues that scientific theories typically go through three stages. When they are first proposed they are regarded as mere algorithms or "black boxes." As they begin to make more successful predictions than already existing theories, they become more respectable "grey boxes." Finally, through their ability to connect previously diverse areas of research, they become standards of intelligibility; they become what Hanson calls "glass boxes." The phenomena described by a theory in this third stage are taken as paradigms of naturalness and comprehensibility. According to Hanson, when a theory has successfully gone through these three stages "our very idea of what 'understanding' means will have grown and changed with the growth and change of the theory. So also will our idea of 'explanation'." [6]

---

[5] *Foresight and Understanding* (New York: Harper & Row, 1963), pp. 45–46.

[6] *The Concept of the Positron* (New York: Cambridge, 1963), p. 38.

This view clearly has a lot of historical support. There are many cases in the history of science where what seems explanatory to one scientist is a mere computational device for another; and there are cases where what is regarded as intelligible changes with tradition.[7] However, it seems to me that it would be desirable, if at all possible, to isolate a common, objective sense of explanation which remains constant throughout the history of science; a sense of 'scientific understanding' on which the theories of Newton, Maxwell, Einstein, and Bohr all produce scientific understanding. It would be desirable to find a concept of explanation according to which what counts as an explanation does not depend on what phenomena one finds particularly natural or self-explanatory. In fact, although there may be good reasons for picking one "ideal of natural order" over another, I cannot see any reason but prejudice for regarding some phenomena as somehow more natural, intelligible, or self-explanatory than others. All phenomena, from the commonest everyday event to the most abstract processes of modern physics, are equally in need of explanation—although it is impossible, of course, that they all be explained at once.

Therefore, although the "intellectual fashion" account may ultimately be the best that we can do, I don't see how it can give us what we are after: an objective and rational sense of 'understanding' according to which scientific explanations give us understanding of the world. We should try every means possible of devising an objective concept of explanation before giving in to something like the "intellectual fashion" account.

From the above discussion of three important contemporary theories of the explanation relation we can extract three desirable properties that a theory of explanation should have:

1. It should be sufficiently general—most, if not all, scientific theories that we all consider to be explanatory should come out as such according to our theory. This is where the "familiarity" theory fails, since, according to that view, theories whose basic phenomena are strange and unfamiliar—e.g., all of contemporary physics—cannot be explanatory. Although it is unreasonable to demand that a philosophical account of explanation should show that every theory that has ever been thought to be explanatory really is explanatory, it must at least square with most of the important, central cases.

---

[7] Examples can be found in Toulmin, Hanson, and T. Mischel, "Pragmatic Aspects of Explanation," *Philosophy of Science*, XXXIII, 1 (March 1966): 40–60.

2. It should be objective—what counts as an explanation should not depend on the idiosyncracies and changing tastes of scientists and historical periods. It should not depend on such nonrational factors as which phenomena one happens to find somehow more natural, intelligible, or self-explanatory than others. This is where the "intellectual fashion" account gives us less than we would like. If there is some objective and rational sense in which scientific theories explain, a philosophical theory of explanation should tell us what it is.

3. Our theory should somehow connect explanation and understanding—it should tell us what kind of understanding scientific explanations provide and how they provide it. This is where D-N theorists have been particularly negligent, although none of our three theories has given a satisfactory account of scientific understanding.

Thus, none of the three theories of explanation we have examined satisfies all our three conditions; none of them has succeeded in isolating a property of the explanation relation which is possessed by most of the clear, central cases of scientific explanation, which is common to the theories of scientists from various historical periods, and which has a demonstrable connection with understanding. Is there such a property?

Consider a typical scientific theory—e.g., the kinetic theory of gases. This theory explains phenomena involving the behavior of gases, such as the fact that gases approximately obey the Boyle-Charles law, by reference to the behavior of the molecules of which gases are composed. For example, we can deduce that any collection of molecules of the sort that gases are, which obeys the laws of mechanics will also approximately obey the Boyle-Charles law. How does this make us understand the behavior of gases? I submit that if this were all the kinetic theory did we would have added nothing to our understanding. We would have simply replaced one brute fact with another. But this is not all the kinetic theory does—it also permits us to derive other phenomena involving the behavior of gases, such as the fact that they obey Graham's law of diffusion and (within certain limits) that they have the specific-heat capacities that they do have, from the same laws of mechanics. The kinetic theory effects a significant *unification* in what we have to accept. Where we once had three independent brute facts—that gases approximately obey the Boyle-Charles law, that they obey Graham's law, and that they have the specific-heat capacities they do have—we now have only one—that molecules obey the laws of mechanics. Furthermore, the kinetic theory also allows us to integrate the behavior of gases with other phenomena, such as the motions of the planets and of falling bodies near the earth. This is because the laws of mechanics also permit us to derive both the fact that planets obey Kepler's laws and the fact that falling bodies obey Galileo's laws. From the fact that *all* bodies obey the laws of mechanics it follows that the planets behave as they do, falling bodies behave as they do, and gases behave as they do. Once again, we have reduced a multiplicity of unexplained, independent phenomena to one. I claim that this is the crucial property of scientific theories we are looking for; this is the essence of scientific explanation—science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given. A world with fewer independent phenomena is, other things equal, more comprehensible than one with more.

Many philosophers have of course noticed the unifying effect of scientific theories to which I have drawn attention; e.g., Hempel in one place writes: "a worthwhile scientific theory explains an empirical law by exhibiting it as one aspect of more comprehensive underlying regularities, which have a variety of other testable aspects as well, i.e., which also imply various other empirical laws. Such a theory thus provides a systematically unified account of many different empirical laws" (144). However, the only writer that I am aware of who has suggested that this unification or reduction in the number of independent phenomena is the essence of explanation in science is William Kneale:

> When we explain a given proposition we show that it follows logically from some other proposition or propositions. But this can scarcely be a complete account of the matter. . . . An explanation must in some sense simplify what we have to accept. Now the explanation of laws by showing that they follow from other laws is a simplification of what we have to accept because it reduces the number of untransparent necessitations we need to assume. . . . What we can achieve . . . is a reduction of the number of independent laws we need to assume for a complete description of nature.[8]

But does this idea really make sense? Can we give a clear meaning to 'reduce the total number of independent phenomena'? Can we make our account a little less sketchy? First of all, I will suppose that we can represent what I have been calling *phenomena*—i.e., general uniformities or patterns of behavior—by law-like *sentences*;

---

[8] *Probability and Induction* (New York: Oxford, 1949), pp. 91–92.

and that instead of speaking of the total number of independent phenomena we can speak of the total number of (logically) independent law-like sentences. Secondly, since what is reduced is the total number of phenomena we have to accept, I will suppose that at any given time there is a set K of *accepted* law-like sentences, a set of laws accepted by the scientific community. Furthermore, I will suppose that the set K is deductively closed in the following sense: if S is a law-like sentence, and $K \vdash S$, then S is a member of K; i.e., K contains all law-like consequences of members of K. Our problem now is to say when a given law-like sentence permits a reduction of the number of independent sentences of K. For an example of what we want to characterize, let K contain the Boyle-Charles law, Graham's law, Galileo's law of free fall, and Kepler's laws, and let S be the conjunction of the laws of mechanics. Intuitively, we think S permits a reduction of the total number of independent sentences of K because we can replace a large number of independent laws by one (or at least by a smaller number); i.e., we can replace the set containing the Boyle-Charles law, Graham's law, Galileo's law, and Kepler's laws by $\{S\}$. The trouble with this intuition is that it is not at all clear what counts as *one* law and what counts as *two*. For example, why haven't we reduced the number of independent laws if we replace the set containing the Boyle-Charles law, Graham's law, etc. by the unit set of their *conjunction*? It won't do to say that this conjunction is really not one law but four since it is logically equivalent to a set of four independent laws. For *any* sentence is equivalent to a set of n sentences for any finite n —e.g., S is equivalent to $\{P, P \supset S\}$, where P is any consequence of S. I think the answer to this difficulty may be the following: although every sentence is equivalent to a set of n independent sentences, it is not the case that every sentence is equivalent to a set of n *independently acceptable* sentences—e.g., the members of the set $\{P, P \supset S\}$ may not be acceptable independently of S; for our only grounds for accepting $P \supset S$, say, might be that it is a consequence of S.

I don't have anything very illuminating to say about what it is for one sentence to be acceptable independently of another. Presumably, it means something like: there are sufficient grounds for accepting one which are not also sufficient grounds for accepting the other. If this is correct, the notion of independent acceptability satisfies the following conditions:

(1) **If $S \vdash Q$, then S is not acceptable independently of Q.**
(2) **If S is acceptable independently of P and $Q \vdash P$, then S is acceptable independently of Q.**

(assuming that sufficient grounds for accepting S are also sufficient for accepting any consequence of S).

Given such a concept of independent acceptability, the notion of 'reducing the number of independent sentences' can be made relatively precise in the following way. Let a *partition* of a sentence S be a set of sentences Γ such that Γ is logically equivalent to S and each S' in Γ is acceptable independently of S. Thus, if S is the conjunction of the Boyle-Charles law, Graham's law, Galileo's law, and Kepler's laws, the set Γ containing the conjuncts is a partition of S. I will say that a sentence S is *K-atomic* if it has no partition; i.e., if there is no pair $\{S_1, S_2\}$ such that S_1 and S_2 are acceptable independently of S and S_1 & S_2 is logically equivalent to S. Thus, the above conjunction is not K-atomic. Let a *K-partition* of a set of sentences Δ be a set Γ of K-atomic sentences which is logically equivalent to Δ (I assume that such a K-partition exists for every set Δ). Let the *K-cardinality* of a set of sentences Δ, K-card (Δ), be inf {card (Γ): Γ a K-partition of Δ}. Thus, if S is the above conjunction, K-card $(\{S\})$ is at least 4. Finally, I will say that S *reduces* the set Δ iff K-card $(\Delta \cup \{S\}) < K$-card (Δ). Thus, if S is the above conjunction and Γ is the set of its conjuncts, S does not reduce Γ.

How can we define *explanation* in terms of these ideas? If S is a candidate for explaining some S' in K, we want to know whether S permits a reduction in the number of independent sentences. I think that the relevant set we want S to reduce is the set of *independently acceptable* consequences of S ($\text{con}_K(S)$). For instance, Newton's laws are a good candidate for explaining Boyle's law, say, because Newton's laws reduce the set of their independently acceptable consequences—the set containing Boyle's law, Graham's law, etc. On the other hand, the *conjunction* of Boyle's law and Graham's law is not a good candidate, since it does not reduce the set of its independently acceptable consequences. This suggests the following definition of explanation between laws:

(D1) S_1 explains S_2 iff $S_2 \in \text{con}_K(S_1)$ and S_1 reduces $\text{con}_K(S_1)$

Actually this definition seems to me to be too strong; for if S_1 explains S_2 and S_3 is some independently acceptable law, then S_1 & S_3 will not explain S_2—since S_1 & S_3 will not reduce $\text{con}_K(S_1 \,\&\, S_3)$. This seems undesirable—why should the conjunction of a completely irrelevant law to a good explanation destroy its explanatory power? So I will weaken (D1) to

(D1′) S_1 explains S_2 iff there exists a partition Γ of S_1 and an $S_i \in \Gamma$ such that $S_2 \in \text{con}_K(S_i)$ and S_i reduces $\text{con}_K(S_i)$.

Thus, if S_1 explains S_2, then so does S_1 & S_3; for $\{S_1, S_3\}$ is a partition of S_1 & S_3, and S_1 reduces $\text{con}_K(S_1)$ by hypothesis.

Note that this definition is not vulnerable to the usual "conjunctive" trivialization of deductive theories of explanation alluded to in footnote 1 above; i.e., the conjunction of two independent laws does not explain its conjuncts. Furthermore, my account shows why such a conjunction cannot be a good explanation. It does not increase our understanding since it does not reduce its independently acceptable consequences.

On the view of explanation that I am proposing, the kind of understanding provided by science is global rather than local. Scientific explanations do not confer intelligibility on individual phenomena by showing them to be somehow natural, necessary, familiar, or inevitable. However, our over-all understanding of the world is increased; our total picture of nature is simplified via a reduction in the number of independent phenomena that we have to accept as ultimate. It seems to me that previous attempts at connecting explanation and understanding have failed through ignoring the global nature of scientific understanding. If one concentrates only on the local aspects of explanation—the phenomenon being explained, the phenomenon doing the explaining, and the relation (deductive or otherwise) between them—one ends up trying to find some special epistemological status—familiarity, naturalness, or being an "ideal of natural order"—for the phenomenon doing the explaining. For how else is understanding conferred on the phenomenon being explained? However, attention to the global aspects of explanation—the relation of the phenomena in question to the total set of accepted phenomena—allows one to dispense with any special epistemological status for the phenomenon doing the explaining. As long as a reduction in the total number of independent phenomena is achieved, the basic phenomena to which all others are reduced can be as strange, unfamiliar, and unnatural as you wish—even as strange as the basic facts of quantum mechanics.

This global view of scientific understanding also, it seems to me, provides the correct answer to the old argument that science is incapable of explaining anything because the basic phenomena to which others are reduced are themselves neither explained nor understood. According to this argument, science merely transfers our puzzlement from one phenomenon to another; it replaces one surprising phenomenon by another equally surprising phenomenon.

The standard answer to this old argument is that phenomena are explained one at a time; that a phenomenon's being itself unexplained does not prevent it from explaining other phenomena in turn. I think this reply is not quite adequate. For the critic of science may legitimately ask how our total understanding of the world is increased by replacing one puzzling phenomenon with another. The answer, as I see it, is that scientific understanding is a global affair. We don't simply replace one phenomenon with another. We replace one phenomenon with a *more comprehensive* phenomenon, and thereby effect a reduction in the total number of accepted phenomena. We thus genuinely increase our understanding of the world.

MICHAEL FRIEDMAN

Harvard University

BOOK REVIEWS

Studies in Inductive Logic and Probability. RUDOLF CARNAP and RICHARD C. JEFFREY, eds. Berkeley: University of California Press, 1971. 264 p. $11.00.

This volume is a collection of articles which Carnap had hoped would appear with others in a non-periodical journal of work in progress toward an adequate system of inductive logic. The central piece is Carnap's "A Basic System of Inductive Logic, Part I"; part II is to appear in a second volume. Thus we will have Carnap's last and most developed views on inductive logic, presented in the systematic fashion of which he was always such a master. I will report at length on Carnap's "Part I" as well as his shorter introductory article, after first touching on the contributions of other authors.

1. Jeffrey's "Probability Measures and Integrals" is intended "to introduce and provide a convenient reference for the basic measure theoretic results used elsewhere in these volumes." The material is well presented, with many helpful examples; but the mathematics involved is quite difficult, and any reader who is not already acquainted with the material must have considerable mathematical acumen in order to absorb it. Thus the article serves best as a review.

Two further articles study the principles of non-negative, positive, and bare relevance. Let M be an attribute and E a non-null proposition (not requiring quantifiers for their expression), and let a and b be individuals not needed for expressing E. Then the prin-