



OXFORD JOURNALS  
OXFORD UNIVERSITY PRESS

## The British Society for the Philosophy of Science

---

Scientific Explanation

Author(s): James Woodward

Source: *The British Journal for the Philosophy of Science*, Vol. 30, No. 1 (Mar., 1979), pp. 41-67

Published by: [Oxford University Press](#) on behalf of [The British Society for the Philosophy of Science](#)

Stable URL: <http://www.jstor.org/stable/687416>

Accessed: 17/11/2014 15:53

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Oxford University Press and The British Society for the Philosophy of Science are collaborating with JSTOR to digitize, preserve and extend access to *The British Journal for the Philosophy of Science*.

<http://www.jstor.org>

## Scientific Explanation

by JAMES WOODWARD

In many philosophical discussions of scientific explanation examples of the following sort are taken as typical or paradigmatic.<sup>1</sup>

(*Ex.* 1) All ravens are black.

This is a raven.

This is black.

(*Ex.* 2) All iron expands when heated.

This is iron which has been heated.

This expands.

(*Ex.* 3) All water boils when heated.

This is water which has been heated.

This boils.

Many people have been struck by the apparent differences between derivations like *Ex.* 1-*Ex.* 3 and the sorts of explanations one might most naturally regard as paradigms of good scientific explanation, explanations of the sort one finds in scientific treatises and textbooks. In this essay I shall attempt to isolate one such difference and to argue that it is of fundamental importance. I shall argue that there is a fundamental difference in the kind of understanding provided by derivations like *Ex.* 1-3 and the kind of understanding provided by a good scientific explanation, and that it is a defect in the standard, Hempelian version of the covering-law model that it is insensitive to this difference. I shall contend that this difference is sufficiently important to warrant one in saying that, in an important sense, derivations like *Ex.* 1-*Ex.* 3 are not scientific explanations at all, and that an acceptable scientific explanation must meet another necessary condition in addition to those standardly imposed by covering-law theorists.

*Received 29 November 1977*

<sup>1</sup> The second and third examples are taken, with slight modifications, from Carnap [1967] and Bergmann [1957] respectively.

I shall begin my discussion by describing in some detail several examples of explanations one might naturally describe as scientific. I shall then explore some of the similarities and differences between these explanations and explanations like (*Ex.* 1-3).

Consider, first, an explanation (*Ex.* 4) of Galileo's law in terms of Newton's laws of motion and the law of gravitation. If we assume that the earth is a sphere and that the only force on a falling body is due to the earth's gravitational attraction, we have from the above laws,

$$F = G \frac{mM}{(R+h)^2} = ma$$

where  $F$  is the force on the falling body,  $m$  its mass,  $h$  its height above the surface of the earth,  $M$  the mass of the earth, and  $R$  the radius of the earth. Since  $R \gg h$ , we can take  $(R+h) \approx R$ . Then dividing through by  $m$ , we get

$$a = G \frac{M}{R^2}.$$

When we substitute numerical values for  $G$ ,  $M$  and  $R$ , we obtain  $g$ , the actual acceleration of an object falling freely above the earth's surface.

The second explanation (*Ex.* 5) I want to consider is the standard explanation given in micro-economic theory for why a monopoly which takes over a formerly competitive industry will raise prices and restrict output.

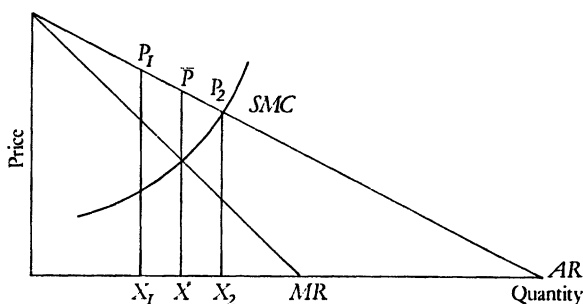


Figure 1

Consider the above diagram. When a monopoly takes over a formerly competitive industry, the demand curve of that industry becomes the demand curve, or the average revenue curve (that is the curve which gives price per unit at each level of output) for the monopolistic firm. This curve, which is labelled  $AR$  in the diagram above, will be downward

sloping—price will be inversely related to quantity of product sold. The curve labelled *MR* in the above diagram is the marginal revenue curve (that is, the curve which gives the change in total revenue occurring with each change in output) for the monopoly. Since the average revenue curve is downward sloping this curve will be more sharply downward sloping. The curve labelled *SMC* is the short-run marginal cost curve for the firm, that is the curve which indicates total change in cost for the monopoly per change in output.

Now it is easy to see that if the monopoly maximises profits it will select that price ( $\bar{P}$ ) at which marginal revenue is equal to marginal cost. Suppose that the firm had selected price  $P_1$ , at which marginal revenue exceeds marginal cost. In that case, if the firm were to lower its price so that the quantity of goods demanded would rise, revenue would increase at a faster rate than costs, so that profits would rise. Thus the profit maximising firm will lower its price toward  $P_1$ . Suppose on the other hand that the firm had selected price  $P_2$ . At this price costs are increasing more quickly than revenue, and profits may be increased by increasing price until  $\bar{P}$  is reached.

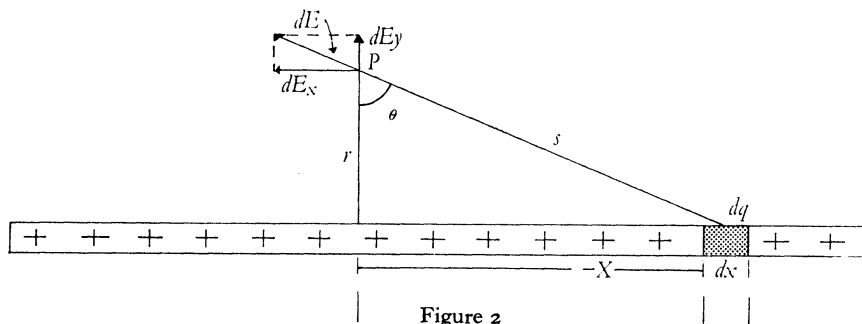
Now contrast the behaviour of the monopoly with the behaviour of a price taking firm in a competitive industry, before it is taken over by the monopoly. Any such firm will sell goods at  $P_2$ , at which price is equal to marginal cost. Such a firm by definition can sell any amount of its output at the going market price. Its average and marginal revenue curve are identical, and it cannot increase profits by restricting the quantity of goods it sells.

Because marginal revenue is always less than average revenue for the monopolistic firm, marginal cost will always exceed marginal revenue at the price  $P_2$  adopted by the competitive firm. Thus the monopoly will always be, in comparison with the price taking firm, a price raiser and output restrictor—it will raise prices from  $P_2$  to  $\bar{P}$  and restrict quantity of goods sold from  $X_2$  to  $\bar{X}$ .

The third explanation (*Ex. 6*) I want to consider is an explanation, in terms of Coulomb's law, of why the magnitude of the electric intensity (force per unit charge) at a perpendicular distance  $r$  from a very long fine wire have a positive charge uniformly distributed along its length, is given by the expression

$$E = \frac{1}{2\pi\epsilon_0} \frac{\lambda}{r}$$

(where  $\lambda$  is the charge per unit length on the wire) and is at right angles to the wire.



We can think of the wire as divided into short segments of length  $dx$ , each of which may be treated as a point charge  $dq$ . The resultant intensity at any point will then be the vector sum of the fields set up by all these point charges. By Coulomb's law, the element  $dq$  will set up a field of magnitude

$$dE = \frac{1}{4\pi\epsilon_0} \frac{dq}{s^2}$$

at a point  $P$  a distance  $s$  from the element. Integrating the  $x$ - and  $y$ -components of  $dE$  separately, we have

$$E_x = \int dE_x = \int dE \sin \theta$$

$$E_y = \int dE_y = \int dE \cos \theta.$$

If  $\lambda$  is the charge per unit length along the wire, we have  $dq = \lambda dx$ , and

$$dE = \frac{1}{4\pi\epsilon_0} \frac{dx}{s^2}.$$

The integration will be simplified if we integrate with respect to  $d\theta$  rather than  $dx$ . From the above figure

$$x = r \tan \theta \quad \text{and} \quad s = r \sec \theta$$

and thus,

$$dx = r \sec^2 \theta \, d\theta.$$

Making these substitutions, we obtain:

$$E_x = \frac{1}{4\pi\epsilon_0} \frac{\lambda}{r} \int \sin \theta \, d\theta$$

$$E_y = \frac{1}{4\pi\epsilon_0} \frac{\lambda}{r} \int \cos \theta \, d\theta.$$

If we assume that the wire is infinitely long, the limits of integration will be from  $\theta = -\pi/2$  to  $\theta = \pi/2$ . Integrating, we obtain

$$E_x = 0$$

$$E_y = \frac{1}{2\pi\epsilon_0} \frac{\lambda}{r}.$$

This shows that the resultant field will be at right angles to the wire and that its intensity is given by

$$E = \frac{I}{2\pi\epsilon_0} \frac{\lambda}{r}.$$

Let me begin my discussion of these three examples by noting that they do indeed conform (or very nearly conform) to the requirements laid down by the proponents of the covering-law model. That is to say, *Ex.* 4–6 are (or are very nearly) sound deductive arguments, arguments in which a law occurs as a premise which is required for the derivation of the explanandum. Moreover, *Ex.* 4–6 seem to meet the other conditions standardly imposed on scientific explanation by covering-law theorists.<sup>1</sup>

It is true that *Ex.* 4 and *Ex.* 6 involve approximations or idealisations which are not, strictly speaking, true, although they are very nearly true. What can be deduced from the laws occurring in *Ex.* 4 and *Ex.* 6 and other true premises are in fact only approximations of the explananda in *Ex.* 4 and *Ex.* 6. *Ex.* 4 and *Ex.* 6 show us why their explananda are very nearly true or why the relationships they describe hold to the extent they do. Nonetheless, it would, I think, be a mistake to attach any great significance to this fact. Certainly it seems to be a mistake to suppose, as some writers apparently have,<sup>2</sup> that satisfying or illuminating scientific explanations always involve approximate rather than strict derivations—*Ex.* 5 is an obvious counter-example and there are many others. It seems to me reasonable to take as our paradigmatic cases of scientific explanation those explanations which are sound deductive arguments and to treat explanations like *Ex.* 4 and *Ex.* 6 as acceptable because they approximate to this pattern.

It also seems clear that, as covering-law theorists have contended, a law or set of laws figures essentially in each of the above derivations (in *Ex.* 4, we have Newton's second law, and the law of gravitation; in *Ex.* 5 we have the law that all firms maximise profits; in *Ex.* 6 we have a version of Coulomb's law).

In what follows I shall assume that the covering-law model is correct as far as it goes; that is, that the covering-law model does state necessary conditions which any acceptable scientific explanation must meet. It is clear, however, that explanations (*Ex.* 1–3) as well as explanations (*Ex.* 4–6) meet all of the necessary conditions on scientific explanation so far discussed. If we wish to develop an account of scientific explanation

<sup>1</sup> I have in mind here the additional requirements imposed by Hempel and Oppenheim in their [1948], supplemented by those imposed by Kim in his [1963], and/or requirements 7.1–7.5 imposed on singular explananda by Raimo Tuomela in chapter VII of his [1973].

<sup>2</sup> A position somewhat like this seems to be taken by Feyerabend in [1962].

which is sensitive to the differences between explanations (*Ex.* 1–3) and explanations (*Ex.* 4–6), we must formulate a further necessary condition on scientific explanation. I suggest the following condition, which I call the requirement of functional interdependence:<sup>1</sup>

- (f) The law occurring in the explanans of a scientific explanation of some explanandum *E* must be stated in terms of variables or parameters variations in the values of which will permit the derivation of other explananda which are appropriately different from *E*.

I shall first illustrate very briefly how this condition applies to the above explanations, and I shall then attempt to clarify it and to motivate its adoption.

Consider the generalisations which figure in explanations (*Ex.* 4–6). These generalisations contain variables or parameters (mass, distance, acceleration, price, quantity of goods, charge, electrical intensity and so forth) which are such that a whole range of different states or conditions can be characterised in terms of variations of their values. The laws in *Ex.* 4–6 formulate a systematic relation between these variables. They show us how a range of different changes in certain of these variables will be linked to changes in others of these variables. In consequence, these generalisations are such that when the variables in them assume one set of values (when we make certain assumptions about boundary and initial conditions) the explananda in the above explanations are derivable, and when the variables in them assume other sets of values, a range of other explananda are derivable. For example, the second law of motion and the law of gravitation which occur in explanation (*Ex.* 4) are such that when the variables in them assume appropriate values (values for the mass and radius of the earth) Galileo's law is derivable. But these generalisations are also such that when the variables in them assume different values (*via* the combination of these generalisations with a different set of initial or boundary conditions) quite different explananda are derivable.

<sup>1</sup> A requirement like the requirement imposed here has been imposed as a requirement on the explanation of scientific laws by several writers. For example, Ernest Nagel [1961], p. 36 holds that

At least one of the premises in the explanation of a given law will meet the following requirements: when conjoined with suitable additional assumptions the premise should be capable of explaining other laws than the given one; on the other hand it should not in turn be possible to explain the premise with the help of the given law even when those additional assumptions are adjoined to the law.

A similar requirement is imposed by Tuomela in [1973], pp. 187 ff. The requirement I impose is considerably stronger than this requirement and is imposed as a requirement on the explanation of singular explananda as well as scientific laws. My discussion below makes clear my reasons for imposing requirement (f) rather than Nagel's requirement.

For example, these generalisations are such that on the assumption that the mass and radius of the earth had different values, a quite different value for the acceleration of a falling body could be derived. These generalisations are also such that we could use them to derive an expression for the rate of fall of a body falling from a distance which is no longer negligible in comparison with the earth's radius. Indeed these generalisations are such that we could use them to derive even more disparate explananda; for example, we could use them in conjunction with other information to derive Kepler's laws and a great many other derivative laws of Newtonian mechanics.

Similarly, the law which occurs in *Ex. 5* can also be used to explain a whole range of different explananda—as *Ex. 5* itself suggests it can be used to explain various features of the behaviour of a firm under competitive conditions and it could be used to explain other features of the behaviour of a monopolistic firm—why it will engage in product differentiation under certain conditions, for example. The law occurring in (*Ex. 5*) is such that it could be used to show that if the initial conditions facing a monopolistic firm were to change in certain ways (assume different values), the behaviour of the firm would change accordingly. (If, for example, the slope of the firm's average revenue curve were to become less steep, it would lower its prices.) And in a similar way the version of Coulomb's law occurring in (*Ex. 6*) can be used, as the parameters in this law assume different values, to explain a range of different explananda—the expressions for electrical intensity along the axis of a uniformly charged ring, or between two equally and oppositely uniformly charged plates, or inside and outside a uniformly charged hollow sphere, for example. By contrast, the generalisations occurring in (*Ex. 1–3*) do not possess these features. For example, 'Raven' and 'black' are not variables which can be used to characterise a range of different values, and 'All Ravens are black' does not formulate a systematic relationship among changes in the values of these variables. The generalisations occurring in (*Ex. 1–3*) are not such that they can be combined with a range of different assumptions about initial conditions to derive a range of different explananda in the way that the generalisations occurring in *Ex. (4–6)* can.

It is this difference between explanations (*Ex. 1–3*) and explanations (*Ex. 4–6*)—the fact that the laws occurring in the former, but not in the latter express a systematic inter-relation between variables which can assume a range of different values, and the fact that the former generalisations but not the latter can, as they assume these different values, be used to derive a range of quite different explananda—which I have attempted to capture by means of the requirement of functional interdependence.



I turn now to the task of explicating and clarifying the notion of functional interdependence. First, when are two explananda 'appropriately different'? The idea I want to capture here is the idea that the generalisations in a successful scientific explanation will permit the derivation of explananda which differ in the way in which, say, the expression for the electrical intensity at a distance from a long, uniformly charged wire differs from the expression for the electrical intensity between two equally and oppositely uniformly charged plates and not merely in the relatively trivial way in which '*a* is black' and '*b* is black' differ. I think that we can capture this idea for singular explananda if we say that two singular explananda *Ba* and *Cb* are appropriately different if and only if  $(x) Bx$  does not entail  $(x) Cx$  and  $(x) Cx$  does not entail  $(x) Bx$ . For non-singular explananda, I shall say that  $E_1$  and  $E_2$  are appropriately different if and only if  $E_1$  does not entail  $E_2$  and vice-versa.

A more difficult problem arises with regard to specifying the meaning of the phrase 'variations in the value of a variable'. First, why do I use this cumbersome and obscure expression at all? Why don't I adopt the following simpler and more straightforward formulation of the requirement of functional interdependence?

(*f'*) The law occurring in the explanans of a scientific explanation of some explanandum *E* must be such that in conjunction with some appropriate set of initial or boundary conditions, it can be used to derive an explanation which is appropriately different from *E*.

Some of my reasons for preferring (*f'*) will only emerge later in this essay. Nonetheless it may be helpful at this point to indicate why a simpler formulation like (*f'*) won't do.

Consider to begin with the 'explanation' (*Ex. 7*)

All ravens are black. All diamonds are green.

*a* is a raven

∴ *a* is black

This explanation meets requirement (*f'*)—it is a generalisation which in conjunction with other initial conditions could be used to derive quite different explananda. Yet it clearly fails to exemplify the pattern we found in explanations (*Ex. 4–6*). If we do not think that (*Ex. 1*) is an acceptable scientific explanation it is difficult to see why the addition of the apparently unrelated generalisation 'All diamonds are green' to the explanans of (*Ex. 1*) should turn (*Ex. 1*) into an acceptable explanation. Clearly we need to require that the generalisation occurring in a scientific explanation be such that it can be used to explain a variety of different explananda in terms of the same parameters or explanatory categories.

But it won't do either to require simply that the generalisation be such that it can be used in conjunction with the same initial condition to derive a range of different explananda. Consider (*Ex.* 8).

All ravens are black. All ravens are cold-blooded.

*a* is a raven

∴ *a* is black

Here both the blackness and cold-bloodedness of a bird is 'explained' by reference to its raven-ness. Nonetheless if we do not think that (*Ex.* 1) and (*Ex.* 7) are scientific explanations it is difficult to see why we should take (*Ex.* 8) to be a scientific explanation. It is hard to see how our understanding of why *a* is black is increased when we are told that all ravens are cold-blooded.

The pattern of explanation achieved in (*Ex.* 4–6) is in fact quite different from that achieved in arguments (*Ex.* 7) and (*Ex.* 8). The parameters appealed to in (*Ex.* 4–6) figure in the derivation of a range of different explananda, but they do so by assuming what might naturally be described as a range of different values. Informally, we can think of the parameters in (*Ex.* 4–6) as associated with scales containing different gradations, the values of these parameters being positions on these scales. The generalisations (*Ex.* 4–6) can then be thought of as showing us how certain movements along these scales (certain changes in the value of these parameters) are systematically associated with movements along other scales (changes in the values of other parameters). For example, the generalisations in *Ex.* 4 are such that they show us how increases or decreases in the mass or radius of the earth, or the distance above the earth's surface from which a stone is dropped are associated with corresponding changes in the explanandum.

It is this feature of the above explanations which I have sought to capture by use of the phrase 'value of a variable'. The sense I want to attach to this expression is roughly this: a variable may be said to have values when we may associate with the variable an ordinal and not merely nominal scale, when it is possible to talk of 'more or less' (in some non-trivial sense) in connection with the values of the variable.

Thus, to begin with, I do not understand the requirement of functional interdependence in such a way that only generalisations containing predicates associated with a ratio scale, like 'mass' or 'length', satisfy the requirement of functional interdependence. Consider, for example, explanations of consumer choice behaviour in terms of the generalisation that consumers will maximise expected utility, where utility values are measured by the method of von Neumann and Morgenstern, which

D

establishes an interval scale. Such explanations do not merely allow us to derive a sentence about how an economic agent will behave in certain circumstances, given his utility schedule and beliefs about the probability of various outcomes. They involve a generalisation which can be used to tell us how, if an economic agent's utility schedule or probability beliefs alter in various ways, his choice behaviour will change accordingly. They are thus explanations which satisfy the requirement of functional interdependence. (They may of course have other infirmities.)

A generalisation may meet the requirement of functional interdependence even if it does not contain predicates which are associated with an interval scale, so long as it contains predicates which are associated with a rough ordinal scale. For example, in his essay, 'Imperfect Rationality' (Watkins [1970]) John Watkins introduces what he calls a 'step-wise likelihood scale' of subjective probability. This is a scale consisting of broad discontinuous gradations which represent a subject's classification of various events as to their subjective probability. The subject may be unable to say whether he thinks a war between the U.S. and Canada is more or less likely than a war between Britain and France, but he may think that both these possibilities fall within a broad gradation which places them considerably lower on the likelihood scale than war between Israel and Syria. And he may think that the likelihood of war between Russia and China occupies some intermediate gradation on the scale. If we assume that the value the subject assigns to various outcomes may be similarly scaled we might use the 'law' that he will maximise expected utility to explain why he chooses as he does under a variety of different kinds of situations involving uncertainty. Here only crude comparison among the subject's beliefs regarding the probabilities of various outcomes and the values he assigns to these different outcomes are possible, and we are not, as in the example above, able to measure these on an interval scale, but nonetheless at least in some cases it seems possible, given the subject's preferences and beliefs, not only to derive how he will choose, but to say how, if his preferences and beliefs had been different, he would have chosen differently. Here too, I want to say that it is appropriate to talk of the 'values' of the subject's beliefs and preferences and of how, if those preferences and beliefs had assumed different values, he would have chosen in various different ways. Here too I want to say that this generalisation may occur in explanations satisfying the requirement of functional interdependence.

Consider another case. In his recently published study, *Religion and Regime* ([1967]) Guy Swanson undertakes to classify political regimes on a scale with five gradations, according to the extent to which govern-

mental power in the regime is shared by the members of the political community. He also attempts to classify various religions according to the extent to which they subscribe to a belief in the immanence of God. He suggests that the items on these two scales are systematically inter-related—in particular that the extent to which governmental power in a regime is concentrated is positively correlated with the degree to which the religion to which it subscribes is committed to a belief in the immanence of God. Thus, for example, centralist regimes with a single autocrat (regimes in which power is most concentrated) are said to be typically Roman Catholic (the religion which most emphasises the immanence of God), while regimes in which a number of groups play an important role in determining governmental policy are said to be typically Calvinistic (the religion which least emphasizes the immanence of God). Regimes in various intermediate positions on the concentration-of-power scale are correlated with religions that take various intermediate positions with respect to divine immanence (Lutheranism, Anglicanism).

I do not wish to defend Swanson's 'theory' (which in fact seems to me to be rather implausible), but rather to draw the reader's attention to the fact that it seems to satisfy, in an admittedly crude way, the requirement of functional interdependence. In Swanson's account, various gradations in two magnitudes—concentration of political power and degree of belief in divine immanence—are distinguished. Swanson does not merely advance the hypothesis that some specific degree of concentration of political power is correlated with some particular religion. He also makes claims about how a range of states of concentration of political power are correlated with a range of values of a certain religious variable. His theory suggests how if the concentration of political power within a regime varies, its religious character will change and in doing so satisfies the requirement of functional interdependence with respect to certain explananda.

Similarly other predicates associated with step-wise scales which establish a crude order—indeed any vocabulary which allows us to talk about a range of phenomena in terms of gradations in a few basic parameters might conceivably occur in a scientific law which satisfies the requirement of functional interdependence.

Those cases in which, on my use of the phrase, it will be inappropriate to talk of the 'values of a variable' are cases which involve what is sometimes called a nominal scale, a scale which does not indicate an order, but only sameness or difference. Suppose, for example, that we say that  $x$  is a gem if it is an emerald, ruby, or beryl. Suppose that we also say that the gem-value of  $x = 1, 2$  or  $3$  when  $x$  is respectively an emerald, a ruby, or a beryl and that the colour-value of  $x = 2, 4$  or  $6$  when  $x$  is respectively

green, red or blue. Consider now the generalisation ( $G_1$ ) which tells us that for any gem, its colour-value will be twice its gem-value. Should we say that 1, 2 and 3 are values of the variable 'gem-value' and that ( $G_1$ ) satisfies the requirement of functional interdependence with regard to ( $A$ ) 'All emeralds are green?' Clearly we want to avoid this conclusion, for ( $G_1$ ) is no more than the conjunction of ( $A$ ) and ( $B$ ) 'All rubies are red' and ( $C$ ) 'All beryls are blue'.

The crucial difference between the artificial predicate 'gem-value' and the other predicates I have been considering seems to be something like this: in explaining what the 'gem-value' of a gem was we simply stipulated what this value was for various gems. The only requirement on this stipulation was that each of the gems we considered be assigned some (one) number, and that the same number not to be assigned to different gems. Because of this the predicate 'gem-value' does not describe an order in any non-trivial sense. If we were to discover other kinds of gems (say, sapphires or rhinestones) and ask what their gem-values were, or how (even very roughly) their gem-values compared with the gem-values of the other gems discussed above, we would have no idea how to go about answering this question, even in principle. We may contrast this with the variables 'subjective probability' or 'governmental centralisation' which Watkins and Swanson proposed to introduce. It is true that it may be difficult to say where many governments or beliefs about likelihood will fall on Swanson's or Watkins's scale, but nonetheless we have some idea about what sorts of considerations are relevant to such classifications. Presumably there are some previously unclassified governments or beliefs that we would know, roughly at least, where to put on such scales, and, we can imagine the addition of further steps or other refinements to such scales. These features of predicates like 'subjective probability' and 'governmental centralisation' reflect the fact that, unlike a predicate like 'gem-value', they amount to something more than, so to speak, simply a list of the values they may take for various arguments. It is only when a variable possesses these features—only when in some nontrivial sense it is associated with an ordered scale—that I shall speak of 'values' of the variable, and it is only when a generalisation is stated in terms of variables which may assume different values that it will be possible for it to meet the requirement of functional interdependence.

I want now, by way of conclusion to this section, to note that the requirement of functional interdependence is at best a necessary condition and not a sufficient condition for an acceptable scientific explanation. There are many generalisations which exhibit the pattern we have called functional interdependence with respect to different potential explananda

and yet which should not be regarded as explaining those explananda. We may, for example, have a theory in which the values of two variables  $U_1$  and  $U_2$  exhibit some regular relationship but in which  $U_1$  and  $U_2$  are thought of as causally unrelated, the regular relationship between them being thought of as explainable in terms of some third set of variables or conditions. In such a case we may not be able to explain variations in the value of  $U_1$  in terms of variations in the value of  $U_2$ , even though  $U_1$  and  $U_2$  may exhibit the pattern of inter-relation we have called functional interdependence. For example, the Franz–Wiedemann law states that  $k/T\sigma$  is a constant, where  $k$  is thermal conductivity,  $\sigma$  is electrical conductivity and  $T$  is absolute temperature. But while this law satisfies the requirement of functional interdependence with respect to a number of different explananda, it is generally not supposed that we might use it, in conjunction with statements about the absolute temperature and thermal conductivity of a given piece of metal, to explain why the metal has the electrical conductivity it does. Rather the electrical and thermal conductivities of the metal, as well as the general relationship between them expressed in the Franz–Wiedemann law are thought of as explainable in terms of other features of the metal.

There is another very closely related feature of scientific explanation that talk of functional interdependence does not capture—it is insensitive to what is sometimes called the ‘direction’ of explanation. The account of scientific explanation presented here has nothing to say about why (to use Sylvain Bromberger’s example) ([1966]) we are inclined to suppose that we can explain the period of a pendulum by reference to its length and yet not inclined to suppose that we can explain its length by reference to its period.

These examples suggest that a fully acceptable model of scientific explanation will need to embody some characteristically causal notions (*e.g.*, some notion of causal priority), or some more generalised analogue of these (*e.g.*, some notion of explanatory priority).<sup>1</sup> I have not, in my

<sup>1</sup> I leave open the question of how these causal features of scientific explanation are ultimately to be analysed or understood. It may well be that these features can be explicated in terms of some model of scientific explanation which does not presuppose them. Evan Jobe’s recent discussion of Bromberger’s pendulum example (Jobe [1976]) seems to proceed along these lines. According to Jobe, we can deductively explain why a pendulum has a certain length without making use of the fact that it has a certain period, but any deductive explanation of its period will necessarily involve the fact that it has a certain length. This difference, according to Jobe, accounts for our willingness to explain a pendulum’s period in terms of its length and our unwillingness to explain its length in terms of its period. Thus in Jobe’s discussion the ‘directional’ character of scientific explanation is not regarded as primitive but is itself something which can be explicated in terms of a more basic notion of scientific explanation. If Jobe’s discussion is correct, it provides additional support for my neglect of the ‘directional’ features of scientific explanation in this essay.



remarks above, attempted to deny this but have rather contended that the covering-law model must also be supplemented in a quite different way, along the lines suggested by the requirement of functional interdependence. The features of a good scientific explanation which I have attempted to capture via the requirement of functional interdependence make a contribution to its explanatory power which is at least in part independent of the contribution made by the causal features of scientific explanation mentioned immediately above. The difference between explanations like *Ex. 2* and *Ex. 3* and an explanation like *Ex. 4* does not consist in the fact that the former explanations are causal and the latter is not. A *D-N* explanation may contain a causal law and yet fall short of being a scientific explanation in the sense that *Ex. 4-Ex. 6* are scientific explanations.

**2** While I have attempted to clarify the requirement of functional interdependence I have so far said relatively little to motivate its adoption. I can best begin to do this by developing the contrast between my account of the way in which a scientific explanation provides understanding and Hempel's account in a bit more detail. In his essay 'Aspects of Scientific Explanation' ([1965], p. 327), Hempel writes<sup>1</sup>

The *D-N* argument shows that, given the particular circumstances and 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.

If we confine ourselves to those cases of scientific explanation which involve deterministic laws we can say, I think, without serious distortion, that for Hempel a scientific explanation explains by exhibiting a nomologically sufficient condition for the explanandum, by showing us that, given certain laws and initial conditions the explanandum-phenomenon 'had' to occur, that it could be expected with certainty.

On my view, in contrast, an adequate scientific explanation provides understanding not merely by showing us this, but also by showing us how, if matters had been different in certain respects, other outcomes besides the explanandum phenomenon would have ensued. A scientific

<sup>1</sup> Cf. also 'Aspects of Scientific Explanation', pp. 367-8 where the following condition is imposed as a 'general condition of adequacy for any rationally acceptable explanation of a particular event':

Any rationally acceptable answer to the question 'Why did event *X* occur?' must offer information which shows that *X* was to be expected—if not definitely, as in the case of *D-N* explanation than at least with reasonable probability. Thus the explanatory information must provide good grounds for believing that *X* did in fact occur; otherwise that information would give us no adequate reason for saying 'That explains it—that does show why *X* occurred'.

explanation not only shows that the explanandum phenomenon was to be expected, but also enables us to answer questions of the form 'What would have happened if . . .'. A successful scientific explanation accomplishes this by exhibiting the explanandum phenomenon as one of a range of states, any one of which might have occurred had initial conditions, boundary conditions, and so forth been different in various ways from what they actually were. We are then shown why, conditions being what they are, the explanandum phenomenon rather than one of these alternative outcomes occurred. In effect we are not just shown that the explanandum phenomenon had to occur, but are given some sense for the range of conditions under which it would have occurred. (It is, I take it, clear enough how, in meeting the requirement of functional interdependence, *Ex.* 4–6 exhibit these features.)

But why should this additional information have explanatory significance? One way to appreciate the significance of this additional information is to note that if we simply require that an explanans provide a nomologically sufficient condition for the explanandum we do not insure that the explanans is relevant to the explanandum. When we require, in addition, that the laws in an explanans be such that they could be used to answer a set of what-if-things-had-been-different questions, we help to insure that the explanans will perspicuously identify those conditions which are relevant to the explanandum being what it is.

Let me begin with a very simple and thus possibly misleading example, which is taken from Wesley Salmon's essay 'Statistical Explanation and Statistical Relevance' [1971]. Consider the generalisation ( $L_9$ ) 'All men who take birth control pills regularly, won't get pregnant'. This generalisation is universal in scope and supports counter-factuals. It seems to satisfy the usual syntactic conditions for law-likeness.<sup>1</sup> Can it then be used to explain why Mr Jones, a man who has been taking birth control pills regularly, fails to get pregnant? That is, is

(*Ex.* 9)

( $L_9$ ) All men who take birth control pills regularly fail to get pregnant.

( $C_9$ ) Mr Jones is a man who takes birth control pills regularly.

( $E_9$ ) Mr Jones fails to get pregnant.

an acceptable explanation?

I think it is clear that (*Ex.* 9) is a defective explanation. While the explanans of this explanation does indeed exhibit a nomologically sufficient condition for the explanandum, it does not identify a set of factors or conditions which are relevant to the explanandum. This is of course

<sup>1</sup> I put aside questions about whether the class of men, as a subclass of a biological species, involves an implicit spatio-temporal reference that disqualifies ( $L_9$ ) from lawlikeness.



reflected in the fact that even if Mr Jones stopped taking birth control pills, he still would not get pregnant. Given that Mr Jones is a man, whether or not he takes birth control pills has, as we say, 'nothing to do' with whether or not he gets pregnant.

Now contrast ( $L_9$ ) with ( $L_{10}$ ):

- ( $L_{10}$ ) All women who meet condition  $K$  ( $K$  has to do with whether the woman is fertile, has been having intercourse regularly and so forth) and who take birth control pills regularly will not get pregnant and furthermore all women who meet condition  $K$  and do not take birth control pills regularly will get pregnant.

Suppose that ( $C_{10}$ ) Mrs Jones is a female who meets condition ( $K$ ) and has been taking birth control pills regularly. Can ( $L_{10}$ ) and ( $C_{10}$ ) be used to explain ( $E_{10}$ ) why Mrs Jones doesn't get pregnant? Here of course we have considerably more inclination to say that at least a crude explanation of  $E_{10}$  has been provided. The reason is obvious—whether or not Mrs Jones is taking birth control pills and meets condition  $K$  does have a lot to do with whether she gets pregnant. Here the explanation given (let us call it *Ex. 10*) identifies and exhibits conditions which are not merely sufficient for, but are also relevant to  $E_{10}$ .

This difference is of course reflected in the fact that the explanation given for Mrs Jones's non-pregnancy shows us, while (*Ex. 9*) does not, how if conditions had been different, a different outcome would have ensued. That is to say,  $L_{10}$  is such that it could be used in conjunction with a statement that Mrs Jones is not taking birth-control pills to derive the result that Mrs Jones does get pregnant. By contrast, even if we attempt to supplement  $L_9$  along the lines of  $L_{10}$ ,  $L_9$  will not have this feature—it cannot be used in connection with some different set of initial conditions to derive some explanandum appropriately different from  $E_9$ . We may say that the explanandum given for Mrs Jones's non-pregnancy does, while explanation (*Ex. 9*) does not, identify in a crude way the range of conditions under which the associated explanandum will hold. In contrast to (*Ex. 9*), (*Ex. 10*) shows us why, a certain condition being what it is, the explanandum of (*Ex. 10*) rather than certain alternatives was realised and in so doing so shows how this condition (Mrs Jones's taking birth control pills) makes a difference for, or is relevant to, the explanandum.

I want to suggest that the features of (*Ex. 4*) which have to do with its satisfaction of the requirement of functional interdependence are important because they play an analogous role in the explanation of Galileo's law, because they help to insure that (*Ex. 4*) identifies conditions

which are relevant to and not merely nomologically sufficient for the truth of Galileo's law. It is when we have an explanation which perspicuously identifies the range of conditions under which Galileo's law holds, which is such that it shows how if various conditions were different in various ways explananda other than Galileo's law would be true, that we have an explanation which shows us how these factors are relevant to the truth of Galileo's law. In a similar way it is because explanation (*Ex.* 5) does not merely provide a sufficient condition for its explanandum—that a firm facing certain conditions will exhibit a certain kind of behaviour—but also employs a generalisation which can be used to show how if the firm were faced by various different conditions (if for example, it did not face a downward-sloping demand curve) it would behave in a variety of different ways, that we can think of (*Ex.* 5) as perspicuously identifying conditions which are not merely sufficient for, but relevant to the behaviour of a monopolistic firm.

In (*Ex.* 4) and (*Ex.* 5) we have of course a considerably more detailed and general identification of the conditions under which the explananda involved obtain than we do in the case of (*Ex.* 10). (This is one of the reasons why (*Ex.* 4) and (*Ex.* 5) are considerably better explanations than (*Ex.* 10), which does not qualify as scientific at all according to our standards, since it does not involve an ordinal scale.) In an explanation involving  $L_{10}$  there are in effect two possible initial conditions which may obtain—Mrs Jones either may or may not take birth control pills—and two possible explananda—Mrs Jones either may or may not get pregnant. In a successful scientific explanation we have a kind of generalised analogue of this feature—the explanation identifies not two but a great range of possible explananda, and a range of possible initial conditions under which these different explananda will be realised. The explanation explains in part in virtue of showing us how it is that it was the explanandum rather than one of these many alternative possibilities that was realised and in doing so, perspicuously identifies those conditions which are relevant to the obtaining of these various explananda.

Consider another case. Suppose that  $C$  is a consumer (or rational agent) who chooses alternative  $A_1$  over alternative  $A_2$ . Suppose that we undertake to explain his choice behaviour by reference to the generalisation: ( $G_2$ ) All consumers (or rational agents) will choose  $A_1$  over  $A_2$ . Now contrast this explanation in terms of ( $G_2$ ) with an explanation of  $C$ 's choice behaviour in terms of an account like that proposed by Watkins. In comparison with the explanation in terms of ( $G_2$ ) the explanation in terms of Watkins's account identifies in a much more perspicuous way those factors or conditions which are actually relevant to  $C$ 's choice. On

the explanation in terms of Watkins's account we are shown that  $A_1$  is chosen because  $A_1$  is the choice that will maximise  $C$ 's expected utility given that  $C$  has the preferences and beliefs regarding the likelihood of various outcomes that he does.

On the explanation in terms of ( $G_2$ ) we learn only that  $C$ 's choosing as he does has something to do with the fact that he is a consumer, or a rational agent. Our sense that the explanation in terms of Watkins's account has more perspicuously identified those factors which are relevant to  $C$ 's choice is reflected, I have been arguing, in the fact that in this case we have an explanation which satisfies the requirement of functional interdependence, an explanation which shows us how, if things had been different in various ways (if  $C$  had a different set of preferences or a different set of probability beliefs),  $C$  would have chosen in a variety of different ways.

Consider the last example. Contrast explanation (*Ex. 1*) above of why some raven,  $a$ , is black with what might more naturally be described as a scientific explanation of this explanandum. A scientific explanation of why some raven is black would, on my view, not simply involve some generalisation like 'All ravens are black', but would rather involve something like this: an identification of those specific biochemical reactions within ravens which produce their distinctive pigmentation, and a specification of the genetic mechanisms which are responsible for those reactions. An explanation of this kind (let us call it *Ex. 11*) would, of course, satisfy the requirement of functional interdependence. It would, for example, presumably involve an appeal to mechanisms which could be used to explain why some non-black but otherwise raven-like bird has the colour it has. It might also be an account which could be used to explain why birds of other related species have the colours they do. It would, in any event, be an account which identified the range of conditions under which the explanandum would hold, an account which makes clear how if the genetic structure of ravens, or their biochemistry were to alter in certain ways, their colour would also change. The fact that the scientific explanation of a raven's blackness possesses these features is, I contend, closely bound up with the fact that it seems to identify those factors (the genetics and biochemistry of the raven) which are relevant to the raven's blackness in a relatively perspicuous way, while explanation (*Ex. 1*) does not.

This example also suggests another important point about scientific explanation. Often our background knowledge will create definite expectations about the range of additional explananda a successful scientific explanation must be able to explain and thus definite expectations about

how the requirement of functional interdependence must be met. Even if all the ravens we have observed are black, we know that there are 'anomalous' members of other species which do not have the colours characteristic of those species. We also know that the colours of many species seem to vary with their geographical location—for example, many organisms which are white in colour are found in regions in which there is a considerable amount of snow. This background knowledge creates the expectation that the account we give of the blackness of ravens will have something in common with the account we give of the colour of certain other species and because we expect that the latter account will say something about the conditions under which anomalously coloured members of those species will occur or under which the colouration of those species will vary with geographical location, we expect that the account we give of the blackness of ravens will do a similar thing. To the extent that the account we give of the blackness of ravens remains isolated and unintegrated with our background knowledge (and note how this is the case with a generalisation like 'All ravens are black') we fail to provide an adequate scientific explanation. Here, too, we see that we may have an account which exhibits a nomologically sufficient condition for the blackness of some raven, and yet which, as a consequence of its inability to explain certain other explananda, fails to provide a scientific explanation for the blackness of that raven.

There is another way of putting the contrast between my account of scientific explanation and Hempel's account. For Hempel scientific explanation is a 'local' affair. The question of whether an explanans  $E_1$  explains an explanandum  $E_2$  is thought of as a question which can be resolved simply by focusing our attention on  $E_1$  and  $E_2$ . That is to say, the relation between  $E_1$  and other explananda which are quite different from  $E_2$  is not thought of as relevant to the question of whether  $E_1$  explains  $E_2$ . Given this sort of orientation it is natural to think of  $E_1$  as explaining  $E_2$  by exhibiting (in the paradigmatic case) a nomologically sufficient condition for  $E_2$ . (Given this focus on the local aspects of explanation, what more could  $E_1$  do?)

By contrast, on my view scientific explanation is a more global or systematic affair. Whether or not  $E_1$  explains  $E_2$  depends in part on the relation between  $E_1$  and other sentences which are quite different from  $E_2$ . A scientific theory does not confer intelligibility on a set of phenomena *via* a series of local, independent exhibitions of those phenomena as necessitated. On my account the kind of understanding provided by a scientific theory rather has to do with the ability of that theory to draw together an apparently disparate set of phenomena, to account

systematically for these phenomena in terms of variations in the values of the same set of parameters in some small set of laws.<sup>1</sup>

Finally, by way of conclusion to this section I want to mention two additional features of the account of scientific explanation I have sketched above. First, we may note that my account enables us to understand why the development of abstract, uniform, homogenous vocabularies generally plays such a crucial role in the construction of scientific theories. If a law or theory is to satisfy the requirement of functional interdependence in any very strong way, we must be able to characterise a large range of states in terms of variations in the values of a few basic parameters. To the extent that the vocabulary of a theory contains sharp qualitative transitions and dichotomies rather than continuous gradations, the theory will be unable to achieve the uniform treatment of a range of cases necessary to satisfy the requirement of functional interdependence. Because the vocabulary of ordinary language is generally a vocabulary which contains sharp qualitative transitions and dichotomies, it is generally not a vocabulary which can be used to formulate generalisations which satisfy the requirement of functional interdependence to any very significant degree. It is largely for this reason that scientific theories are not stated in the vocabulary of ordinary language and that the construction of a scientific theory generally requires a switch to quite different vocabulary. The characteristic flattening out or homogenising of the world which is achieved through the use of words like 'mass', 'velocity', 'energy', 'utility', 'force', and so forth is not a fortuitous feature of scientific theories but rather makes an essential contribution to their explanatory power.<sup>2</sup>

A similar set of remarks can be made about the role played by the development of systems of measurement in the construction of scientific theories. It has often been noted that the development of precise systems of measurement is one of the most distinctive features of modern science, and it is commonly held that the scientific status of a discipline is in some way closely bound up with the availability of appropriate techniques of measurement for that discipline.

<sup>1</sup> The terms 'local' and 'global' as well as the associated contrast between a conception of explanation in which scientific explanations 'confer intelligibility on individual phenomena by showing them to be . . . necessary' and a conception of explanation which stresses the systematic features of scientific understanding are taken from Michael Friedman's essay 'Explanation and Scientific Understanding' [1974]. However the model of scientific explanation Friedman develops differs in a number of respects from my own.

<sup>2</sup> These remarks connect up with, and provide a further rationale for Donald Davidson's claim that we may typically expect the law 'underlying' a singular causal sentence to be stated in a technical, non-ordinary vocabulary which differs from the ordinary vocabulary in which the singular causal sentence is stated (*cf.* Donald Davidson [1967]). This point is explored in more detail in my 'Singular Causal Explanation in History' (*forthcoming*).

My account of scientific explanation enables us to understand why measurement should have this sort of significance—it links the explanatory power of a scientific theory to the availability of such techniques of measurement. By contrast, while there are many reasons that have nothing to do with explanation which a Hempelian might have for preferring a theory which contains quantitative laws or in which precise measurements are employed, it is not easy to see how the presence of these features might confer an explanatory advantage on a theory, given a Hempelian conception of explanation.

**3** An account of scientific explanation in terms of functional interdependence also fits naturally with the intuitively attractive idea that a successful scientific explanation does not merely show that the explanandum phenomenon occurs regularly but also exhibits the explanandum phenomenon in a new light, allowing one to see the relevance of certain considerations which were not apparent from the original characterisation of the explanandum. Consider someone who asks for a scientific explanation of why some raven is black. Such a person (let us call him *Q*) does not, I think, merely wish for a demonstration *via* a law and some statement of initial conditions, that this raven ‘had’ to be black. Rather, when *Q* asks for a scientific explanation of why some raven is black, he wishes to know what it is about that raven which makes it black. *Q* is puzzled because he is unable to identify those features of a raven which are relevant to its blackness or is unaware of the laws governing those features. When *Q* is in such a situation he will not be helped, it seems to me, by being told that all ravens are black. This generalisation simply tells *Q* that all other ravens have the feature he finds puzzling about this particular raven. This information may be of interest to *Q*, but it will serve to generalise rather than dispel his original puzzlement. If *Q* does not know what it is about the explanandum phenomenon which results in its behaving as it does, this puzzlement will not be relieved when *Q* is told that the explanandum phenomenon always behaves in the way *Q* finds puzzling.

Instead, *Q* requires something quite different—an explanation that will draw his attention to further considerations the relevance of which is not apparent from *Q*’s original characterisation of the explanandum under investigation. This is what the scientific explanation of why some raven is black does. It allows us to see the raven not simply as a raven but rather as a system having a certain genetic and biochemical structure and allows us to see the relevance of the laws governing this structure to the raven’s colour. We are shown that the bird possesses one kind of



genetic and biochemical structure among the many different kinds of such structures different organisms may possess, and that in each case it is the character of this structure which is relevant to an organism's colour.

I want to suggest that this is a typical feature of a successful scientific explanation—that typically such an explanation will do more than merely show us that the puzzling explanandum phenomenon occurs regularly, that it will allow us to see the explanandum phenomenon in a new light by drawing our attention to a previously unappreciated set of considerations. I shall express this by saying that a scientific explanation typically involves a 'reconstrual' of the explanandum. It should be clear that an explanation which meets the requirement of functional interdependence will, at least in many cases, possess these features. This is because an explanation which meets the requirement of functional interdependence involves seeing the explanandum phenomenon or state as one instance of a more generally characterised range of phenomena or states, the occurrence of any one of which could be explained in terms of the same generalisation, given the occurrence of the appropriate initial conditions. An explanation which fails to meet the requirement of functional interdependence—an explanation like (*Ex. 1*)—will also fail to exhibit its explanandum in a new light or to introduce a set of considerations the relevance of which was previously unappreciated. It is this failure which accounts, I believe, for the air of triviality which surrounds an explanation like (*Ex. 1*). In imposing the requirement of functional interdependence, we require that the generalisation used to explain why some raven is black go beyond this explanandum in some more ambitious sense than that evinced in (*Ex. 1*), that it do something more than simply assure us that the explanandum phenomenon occurs regularly.

Consider another example. Even in the relatively primitive account of choice-behaviour devised by Watkins, a consumer's choice of  $A_1$  over  $A_2$  is not explained by reference to a generalisation like 'all consumers will choose  $A_1$  over  $A_2$ ' but rather by reference to a generalisation to the effect that consumers will maximise expected utility. We come to understand why this particular consumer has chosen  $A_1$  over  $A_2$  not when we are told they all do, but rather when we come to see this choice as an instance of expected utility maximising behaviour. And this latter 'scientific' explanation involves the introduction of considerations which were not apparent from the initial characterisation of the explanandum as a consumer's choice of  $A_1$  over  $A_2$ . The scientific explanation identifies, in virtue of satisfying the requirement of functional interdependence, those factors (the consumer's beliefs and preferences) which are relevant to the consumer's choice, factors which are not identified in any explanation

which simply depends on the generalisation 'all consumers will choose  $A_1$  over  $A_2$ '. An important part of the scientific explanation of the consumer's choice consists in coming to see that these factors make a difference to the consumer's choice—in coming to see that if the consumer's preferences and probability beliefs had been different in various ways he would have chosen differently. It is because an explanation in terms of Watkins's theory satisfies, at least in a crude way, the requirement of functional interdependence that it provides this sort of information and satisfies our demand for a reconstrual of the explanandum.

There is a tendency among many covering-law theorists to think of the explanation of particular facts (*i.e.*, explanations in which the explananda are singular sentences) as a primary or at least typical scientific activity. But, in fact, as the examples of scientific explanation we have discussed suggest, scientific explanations typically have as their explananda generalisations rather than singular sentences. One finds, in scientific treatises and textbooks, explanations of why simple pendula have periods  $T = 2\pi\sqrt{l/g}$ , of why monopolies are output restrictors, of Boyle's and Charles's Law, of Bernouilli's equation, of the expression relating pressure and volume in gasses undergoing an adiabatic process. One does not find in addition to these explanations a distinct kind of explanation in which, *e.g.*, the period of some particular simple pendulum is explained in terms of the generalisation, 'All simple pendula have periods  $= 2\pi\sqrt{l/g}$ ', or in which the output restricting behaviour of some monopolistic firm is explained by reference to the generalisation, 'All monopolistic firms are output restrictors'. To the extent that it makes sense to speak of the scientific explanation of such particular cases, the generalisations which would figure in such explanations are surely just those that would figure in the explanation of the corresponding general case. That is to say, a scientific explanation of why this  $A$  is  $B$  will involve just those generalisations which could be used to explain why all  $A$ s are  $B$ s. In this sense, the scientific explanation of particular facts is an activity which is derivative or parasitic on the scientific explanation of generalisations. The tendency to suppose otherwise plays, I think, no small role in making the picture of scientific explanation covering-law theorists have given us, seem plausible and attractive. Correspondingly, to the extent we see the explanation of particular facts as a derivative scientific activity we will, I think, be led to see the picture of scientific explanation covering-law theorists have given us as misleading in certain crucial respects.

4 Consider one last example, which will help to draw together the disparate strands of my discussion. In a well-known passage in 'Aspects



of Scientific Explanation', Hempel describes two different explanations which I have distinguished by means of brackets for later reference, of why the water level of a beaker containing a piece of floating ice will remain unchanged as the ice melts.

A purely logical point should be noted here, however. If an explanation is of the form  $(D-N)$ , then the laws  $L_1, L_2, \dots, L_r$  invoked in its explanans logically imply a law  $L^*$  which by itself would suffice to explain the explanandum event by reference to the particular conditions noted in the sentences  $C_1, C_2, \dots, C_k$ . This law  $L^*$  is to the effect that whenever conditions of the kind described in the sentences  $C_1, C_2, \dots$ , the explanandum-sentences occurs. Consider an example: A chunk of ice floats in a large beaker of water at room temperature. Since the ice extends above the surface, one might expect the water level to rise as the ice melts; actually, it remains unchanged. Briefly, this can be explained as follows: [According to Archimedes' principle, a solid body floating in a liquid displaces a volume of liquid that has the same weight as the water displaced by its submerged portion. Since the melting does not change the weight, the ice turns into a mass of water of the same weight, and hence also of the same volume, as the water initially displaced by its submerged portion; consequently, the water level remains unchanged. The laws on which this account is based include Archimedes' principle, a law concerning the melting of ice at room temperature; the principle of the conservation of mass, and so on.] None of these laws mentions the particular glass of water or the particular piece of ice with which the explanation is concerned. Hence the laws imply not only that as this particular piece of ice melts in this particular glass, the water level remains unchanged, but rather the general statement  $L^*$  that [under the same *kind* of circumstance, *i.e.*, when any piece of ice floats in water in any glass at room temperature, the same *kind* of phenomenon will occur, *i.e.*, the water level will remain unchanged. . . clearly,  $L^*$  in conjunction with  $C_1, C_2, \dots C_k$  logically implies  $E$  and could indeed be used to explain, in this context, the event described by  $E$ .] ([1965], p. 347)

Whether 'minimal covering-law's like  $L^*$  will always be, in any natural or interesting sense of the word, 'laws' is an interesting question but one which will not detain us here. The questions I want to consider here is rather this: assuming that  $L^*$  is a law, it is a law which could be used in a scientific explanation of why ( $E$ ) the water level in some particular glass remained unchanged when the ice in it melted?

In the above passage from Hempel, I have introduced two sets of brackets. The first set encloses a sketch of what might reasonably or naturally be regarded as a scientific explanation of  $E$ . The second encloses a minimal covering-law explanation of  $E$ . The contrast between these two explanations is, I think, quite striking and much of my discussion in this chapter can be regarded as an attempt to explicate this contrast.

Note to begin with the contrast between  $L^*$  and the scientific laws—Archimedes' principle, the law of conservation of matter and so forth—

which figure in the explanation enclosed in the first set of brackets. The latter but not the former are stated in that abstract vocabulary which makes possible the characterisation of a range of different cases and their systematic inter-relation. The laws which figure in the explanation enclosed in the first set of brackets are such that they can be used in conjunction with different sets of initial conditions to explain a whole range or spectrum of different explananda—thus Archimedes' principle can be used, for example, to explain why any solid body floating in a liquid displaces the volume of liquid it does. The explanation enclosed in the first set of brackets gives us some sense for the range of conditions under which the explanandum will hold. When we see the relevance of the considerations invoked in the explanans—that a solid which floats displaces a volume of liquid equal to its own weight, that no change of weight is involved when a solid melts—we see that the same reasoning could well be used to show that any solid floating in its liquid form and melting will leave the level of the liquid unchanged. Nothing in the explanation of *E* turns essentially on the fact that a piece of ice is involved in the above example. We can also see, once we appreciate the relevance of Archimedes' principle to this case, that it does make a difference whether the solid is floating in the liquid. If, for example, the solid sinks in the liquid, then it will displace a volume of liquid equal to its own volume rather than its weight, and since a given mass of solid will generally increase in volume when it melts, the level of liquid in the container will rise as the submerged solid melts.

By contrast, the minimal covering-law explanation in terms of  $L^*$  gives us no information of this kind. It can be regarded as assuring us, perhaps, that the water level in the glass 'had' to remain unchanged when the ice melted, but it does not answer the 'what if . . .' questions that the explanation enclosed in the first set of brackets can be used to answer. And to say this is to say that the minimal covering-law explanation does not perspicuously exhibit the factors or conditions which are relevant to the water level in the glass remaining unchanged in the way that the explanation enclosed in the first set of brackets does. One can be aware of the information contained in the minimal covering-law explanation, and yet largely fail to understand what it is about an ice cube melting in a glass that results in the water level in the glass remaining unchanged. One may be aware of  $L^*$  and yet still be unclear whether the fact that the water level in the glass remained unchanged has to do with some feature peculiar to water and ice, whether it has to do with the fact that a glass rather than some other container was employed, and so forth. When we ask for a scientific explanation of *E* we are, I think, interested in the

answers to questions like these, not merely in a certification that *E* had to be true.

We may note also the important role that a recharacterisation or reconstrual of the explanandum plays in the explanation enclosed in the first set of brackets and the complete absence of this feature in the minimal covering-law explanation. In the first explanation we begin to understand why the water-level of the glass remained unchanged as the ice-cube melted in it when we come to see the ice cube as simply one instance of a solid floating in a liquid (and hence, according to Archimedes's principle a solid which displaces a volume of liquid equal to its own weight), the melting of the cube as a process, which like any other change of state, results in no weight change, and so forth. The air of triviality which, by contrast, surrounds the minimal covering-law explanation, reflects its failure to provide any such reconstrual of *E*, any such exhibition of *E* in a new light.

These dissimilarities between the two explanations considered above seems to me to warrant my claim that the difference between them may profitably be regarded as a difference in kind rather than merely a difference in degree. The explanation enclosed in the first set of brackets answers a kind of question which is not answered by the explanation in the second set of brackets. The first explanation does something more than merely exhibit a nomologically sufficient condition for *E*, and this something more makes a crucial difference to its character as an explanation. It is this difference I have tried to capture in characterising the first explanation as scientific and the second explanation as non-scientific.

*University of Wisconsin—La Crosse*

#### REFERENCES

- BERGMANN, G. [1957]: *Philosophy of Science*. Madison: University of Wisconsin Press.
- BROMBERGER, S. [1966]: 'Why Questions' in R. G. Colodny (ed.): *University of Pittsburgh Series in the Philosophy of Science*, Vol. III: Pittsburgh, University of Pittsburgh Press, 1966, pp. 86–111.
- CARNAP, R. [1967]: *Philosophical Foundations of Physics*, edited by Martin Gardiner, New York: Basic Books.
- DAVIDSON, D. [1967]: 'Causal Relations', *Journal of Philosophy*, **64**, pp. 691–703.
- FEYERABEND, P. [1962]: 'Explanation, Reduction, and Empiricism' in H. Feigl and G. Maxwell (eds.): *Minnesota Studies in the Philosophy of Science*, vol. III: Minneapolis: University of Minnesota Press, 1962.
- FRIEDMAN, Michael [1974]: 'Explanation and Scientific Understanding', *Journal of Philosophy*, **71**, pp. 5–19.
- HEMPEL, C. G. [1965]: 'Aspects of Scientific Explanation' in Hempel, *Aspects of Scientific Explanation*, New York: The Free Press, 1965.
- HEMPEL, C. G. and P. OPPENHEIM [1948]: 'Studies in the Logic of Explanation', *Philosophy of Science*, **15**, pp. 135–75, reprinted in Hempel, *Aspects of Scientific Explanation*, New York: The Free Press, 1965.

- JOBE, E. [1976]: 'Discussion: A Puzzle Concerning D-N Explanation', *Philosophy of Science*, **43**, pp. 542-9.
- KIM, J. [1963]: 'Discussion: On the Logical Conditions of Explanation', *Philosophy of Science*, **30**, pp. 280-91.
- NAGEL, E. [1961]: *The Structure of Science*. New York: Harcourt, Brace, and World, Inc.
- SALMON, W. [1971]: 'Statistical Explanation' in Salmon, *et al.*, *Statistical Explanation and Statistical Relevance*: Pittsburgh, University of Pittsburgh Press, 1971.
- SWANSON, G. [1967]: *Religion and Regime*. Ann Arbor: University of Michigan Press.
- TUOMELA, R. [1973]: *Theoretical Concepts*. Wien: Springer-Verlag.
- WATKINS, J. [1970]: 'Imperfect Rationality' in Borger and Cioffi (eds.): *Explanation in the Behavioral Sciences*. Cambridge: Cambridge University Press, 1970.