6 Models of Explanation

Introduction (p.675)

Just as the definiendum and definiens constitute what is to be defined and its definition, respectively, one must consider the 'explanandum' and 'explanans': what is to be explained and its explanation. Hempel used these concepts in his covering law thesis, "...according to which genuine explanations are arguments that must contain empirical laws" (p.675). The two principal covering law models proposed were the Deductive-Nomological (D-N) model in 1948, and the Inductive-Statistical (I-S) model in 1962.

Carnap (p.678)

[1966 paper on Hempel's motivations for introducing the covering law models.]

Carnap begins by contrasting his early writings within the Vienna circle, largely reactions against the 19th century German idealists and later the biologist and philosopher Hans Driesch [1867-1941], with his relatively peaceful time as a celebrated philosopher in the United States. A candidate for quote of the week: "...the nature of one's opponents often determines the way in which one's views are expressed" (p.679).

Scientific laws banish metaphysical agents, such as 'the demons behind the cloud that poured rain over us'. Driesch classified living organisms and gave each class its own metaphysical agent: its 'entelechy'. [Note>To Aristotle, who first used the term, an entelechy was a 'realization' or an 'actuality', as opposed to a 'potentiality'.] Driesch referred to a force that causes living things to behave in the way they do. He likened it to a physical unobserved force like magnetism. Carnap and his friends in the Vienna circle were unhappy with this analogy, since a physicist would never explain a nail's movement towards magnetized iron by inventing the word 'magnetism', but would instead explain it through various laws. Some of these laws may well be qualitative. But Driesch did not speak of laws at all, something which made Carnap rate it to be without explanatory power: it could only produce 'pseudoexplanations'. To Carnap, every explanation contains at least one law.

Scientific laws can be used to explain observables. They can also be used for predicting new unobserved facts and phenomena. The unknown does not have to reside in the future, but always has the same logical form> From a universal implication and an instance of its antecedent, infer an instance of its consequent. In the language of logic>

$\forall x(P(x) \rightarrow Q(x))$

$P(a)$

$Q(a)$

As in: All cats are grey. Pixie is a cat. Therefore Pixie is grey.

If you do not have access to a universal, a statistical implication is sometimes used, but then the unknown can only be probably true, not absolutely certain.

As in: Most cats are grey. Pixie is a cat. Therefore Pixie is most likely grey.

The logic of probability then replaces deduction. In everyday life, we use these schemata all the time, although we may well be ignorant about their logical form.

Hempel (p.685)

[A 1962 book chapter.]

A 'deductive-nomological explanation' has two sets of premises, particular facts and general laws, which together form the 'explanans'. The conclusion, the 'explanandum-statement', is a description of an 'explanandum-event'. [Note> When we say explanandum, we may refer either to the 'explanandum-statement' or to the corresponding 'explanandum-event'.] The Greek word 'nomos' means law, and because the general laws must cover the explanandum, the explanation model that uses deductive-nomological explanations is referred to as the 'covering law model'. The 'why?'-questions scientists ask may be transferred also to the meta-level, as in "Why do Galileo's and Kepler's laws hold?" which "is answered by showing that these laws are but special consequences of the Newtonian laws of motion and of gravitation; and these, in turn, may be explained by subsumption under the more comprehensive general theory of relativity" (p.686). In general, causal explanations of all kinds are deductive-nomological in character, Kepler claims [and he also notes that the converse is false].

Following Carnap, Hempel distinguishes between the 'strictly universal' and the 'probabilistic-statistical' form of deductive-nomological explanations. [He uses two horizontal lines for the latter form, and not just one, to graphically separate explanans from explanandum.] Naturally, the latter form produce inductive arguments, and the former deductive ones. "...the two types of explanation as characterized above constitute ideal types or theoretical idealizations and are not intended to reflect the manner in which working scientists actually formulate their explanatory accounts. Rather, they are meant to provide explications, or rational reconstructions, or theoretical models, of certain modes of scientific explanation. In this respect our models might be compared to the concept of mathematical proof (within a given theory) as construed in meta-mathematics" (p.691). Note> Hempel is here referring to work by Gödel on explicating 'follows logically from', and the like.

Hempel again takes inspiration from Carnap when he distinguishes between idealized reasoning and everyday reasoning: "...explanations put forward in everyday discourse and also in scientific contexts are often *elliptically formulated*" (p.691). This means basically that most people are forgiving when it comes to understanding someone else's reasoning; not everything needs to be explained in detail everywhere, some laws or facts could be filled in if necessary. Another way in which this forgiveness can manifest itself is in 'partial explanations'. Hempel gives an example of Freud's slipping of the pen [wishful thinking]. If Freud's argument had been deductive, it would have had to be complete, but Freud's point does come across even though the explanans is not complete. To complete the explanandum is in general not possible “for any particular event may be regarded as having infinitely many different aspects or characteristics, which cannot all be accounted for by a finite set, however large, of explanatory sentences” (p.693). Hempel also names even weaker explanatory accounts as ‘explanation sketches’. Finally, he notes that classifying various kinds of explanations does not lend itself to any objective procedure.

[The conditions of adequacy for deductive-nomological explanations are presented in the commentary, p.770.]

Hempel (p.695)

[Excerpt from his 1965 collection of papers “Aspects of Scientific Explanation”.]

A deductive-nomological explanation is a potential deductive-nomological prediction: this is determined by the timing, as pointed out by Hempel and Oppenheim already in 1948, and named the ‘thesis of the structural identity (or symmetry) of explanation and prediction’. Hempel also believes his symmetry thesis applies to inductive-statistical explanations. The thesis has been criticized. Scheffler [in 1958] looked at the syntactic form of the explanandum and noted that it is a sentence-token> “a concrete utterance or inscription of a sentence purporting to describe some event that is to occur after the production of the token” (p.695). Hempel stresses that the thesis is to be understood as being about arguments, not sentence-tokens, and that it amounts to “the conjunction of *two sub-theses*, namely (i) that *every adequate explanation is potentially a prediction* ...; (ii) that conversely *every adequate prediction is potentially an explanation*” (p.695-6). Hempel defends the first as “an almost trivial truth” in deductive-nomological explanations, since the “explanans logically implies the explanandum” (p.696).

Hempel goes on to define a ‘condition of adequacy’> producing adequate reasons for citing something as an explanation of an event. There are, however, problems with respect to (i) above if this condition is applied to general laws, since laws “purport to express timeless uniformities and thus make no reference to any particular time, whether past, present, or future” (p.696). Thus, adequate explanations of general laws can never be predictions. The condition is necessary but not sufficient for adequate explanations, consider for example statistical explanations [like the lung cancer correlation with smoking].

An argument against (i), formulated by Toulmin is that Darwin’s theory is often credited with great explanatory power but has still not been used for purposes of prediction (of new species, for instance). Hempel felt that such criticism mixed up the story of evolution (which is a narrative explaining nothing) with the theory of evolution. Another kind of argument against (i) has been given by Scriven, who describes a jealous husband who killed his wife> “The fact that the man was jealous might well have been ascertainable before the deed, but to explain the latter, we need to know that his jealousy was intense enough to drive him to murder; and this we can know only after the deed has actually been committed. Here then, the occurrence of the explanandum event provides the only grounds we have for asserting one important part of the explanans; the explanandum event could not therefore have been predicted by means of the explanatory argument” (p.698). Hempel is impressed of the methodological aspects of Scriven’s example, and calls the explanation in this case ‘self-evidencing’. Hempel, however, concludes that such examples in no way threaten the conditional thesis, since Scriven shows only that “sometimes we do not know independently of the occurrence of the explanandum event that all the conditions listed in the explanans are realized. However, this means only that in such cases our conditional thesis is counterfactual, i.e., that its if-clause is not satisfied, but not that the thesis itself is false" (p.699).

Scriven has also protested against (ii), as has Scheffler> "...a list of the results obtained in a long series of tossings of a given coin may provide a good basis for predicting the percentage of Heads and Tails to be expected in the next 1000 tossings of the same coin; but ... that list of data provides no explanation for the subsequent results. Cases like these raise the question of whether there are not sound modes of scientific prediction that proceed from particulars to particulars without benefit of general laws such as seem to be required for any adequate explanation" (p.703).

Hempel (p.706)

[Another excerpt from his 1965 collection of papers "Aspects of Scientific Explanation".]

If we are told a recovered patient was given penicillin following a streptococcus infection, we might consider this an explanation of the patient's recovery. But we cannot invoke any general law to this end. The statement describing the recovery can be inductively inferred from statistical hypotheses, and this only relative to the body of evidence described by these hypotheses. Hempel describes this schematically by an analogy to the distinction between true sentences and necessary (or necessarily true) sentences. A classical syllogism is>

All F are G. a is F. Therefore a is G.

But even if "a is G" is true, it is true only relative to the body of evidence described by the two premises "All F are G" and "a is F". If it were necessarily true, "a is G" would be true as a consequence of its logical form, without reference to the premises. The simplest example would be "a is a". Only slightly more complicated would be "a is a or b".

Note> In modal logic, a statement can be true or false, as usual, but it can also be necessarily true or possibly true. The modalities of 'necessary' and 'possible' are primitives which cannot be removed or re-interpreted in any other terms. There are modal laws, however, governing their inter-relationship. Examples are>

X is possible iff it is not true that it is necessary that X is not true.

X is necessary iff it is not true that it is possible that X is not true.

If X is necessary, then X is possible.

It is not necessary that if X is possible, then X is necessary.

The logic of necessary and possible truths date back to early 20th century, to C. I. Lewis, and is called 'alethic modal logic'. There are other modal logics, like the later 'deontic modal logic' which deals with statements that ought to be true, should be true, etc.

In some modal logics, and in inductive logic [Note> these are both non-classical logics], a phenomenon called 'ambiguity of inductive-statistical explanation' sometimes occurs. Hempel describes this as follows. "...for a proposed probabilistic explanation with true explanans which confers near certainty upon a particular event, there will often exist a rival argument of the same probabilistic form and with equally true premises which confers near certainty upon the nonoccurrence of the same event. And any statistical explanation for the occurrence of an event must seem suspect if there is the possibility of a logically and empirically equally sound probabilistic account for its nonoccurrence. *This predicament has no analogue in the case of deductive explanation*" [Example> Stanford weather] (p.709). If the explanation refers to all truths known to science rather than to the explanans at hand, the same predicament can occur and is then called 'the epistemic ambiguity of statistical explanation' [Note> episteme comes from the Greek word for knowledge]. Since the set of all truths in science contains only true statements, it is logically consistent [i.e. it has a model] and so only one out of two mutually inconsistent conclusions of deductive arguments can become part of that set. Thus, classical logic again escapes the predicament.

Carnap in 1950 formulated what Hempel refers to as a “maxim for the application of inductive logic” (p.711)> the ‘requirement of total evidence’ as: “in the application of inductive logic to a given knowledge situation, the total evidence available must be taken as basis for determining the degree of confirmation. Using only a part of the total evidence is permissible if the balance of the evidence is irrelevant to the inductive ‘conclusion’” (p.711). We might perhaps use this maxim to position our explanans parts of our explanations between the two extremes of using ALL the information available now, and using ONLY the available information. In the former case, all our explanations would have the same explanans; the collected knowledge in science. In the latter case, we risk including information that will later be kicked out from the set of scientific knowledge, because of completion of further tests, for instance. Hempel’s suggests the ‘requirement of maximal specificity for inductive-statistical explanations’ [for an explanation, see p.714].

The requirement disposes of the problem of epistemic ambiguity, Hempel claims, since two rival statistical arguments with high associated probabilities, and with an explanans belonging to the set of scientific knowledge K, at least one of them violates the requirement. “*...the concept of statistical explanation for particular events is essentially relative to a given knowledge situation as represented by a class K of accepted statements*. Indeed, the requirement of maximal specificity makes explicit and unavoidable reference to such a class, and it thus serves to characterize the concept of ‘I-S explanation relative to the knowledge situation represented by K.’ We will refer to this characteristic as the *epistemic relativity of statistical explanation*” (p.715). Finally, the concept of ‘potential statistical explanation’ stipulates that “no matter how much evidential support there may be for the explanans, a proposed I-S explanation is not acceptable if its potential explanatory force with respect to the specified explanandum is vitiated by statistical laws which are included in K but not in the explanans, and which might permit the production of rival statistical arguments” (p.716).

Ruben (p.720)

[Extracts from his book “Explaining Explanation” (1990) in which Ruben discusses various shortcomings of Hempel’s schemata of explanation, the first of which is irrelevance; see Lyon example (p.720), Achinstein example (p.721), and Salmon example (p.724).]

In the Achinstein example, irrelevance is manifested by ‘causal pre-emption’> “Suppose some event, e, has two potential causes c and d, in the sense that c occurs and causes e, and that d also occurs and does not cause e, but would have caused e if c had not occurred. d is a potential alternative cause of e, but is pre-empted by the actual cause c” (p.722). Hempel was here defended by Redhead, who claimed Achinstein had not taken all relevant information into account in the explanans, but Ruben is not impressed> “It is *true* that whoever eats a pound of arsenic at t dies within twenty-four hours, even when sometimes death of arsenic ingestors is actually brought about by buses or something else” (p.723) [and another contender for “Quote of the week”].

[Salmon’s example is neatly illustrated by the 13th century Sufi jester *Nasruddin Mulla*’s tiger seeds. Sufi is easily the most important reconciler of religious doctrines in history.]

Ruben then turns to explanatory symmetries, the other counterexample towards Hempel’s schemata> “These equations or biconditionals will allow the *derivation* of the height of the flagpole from the length of the shadow and the length of the shadow from the height of the flagpole ... But ... the first derivation would be nonexplanatory; the second, explanatory" (p.726) [The size of Ulysses' bedroom example].

Ruben then introduces a 'cure': the 'causal condition', resting on earlier thinkers including Aristotle and Mill> "Given the angle of the sun's elevation, it is the height of the flagpole that causes the length of the shadow, and not vice versa ... So causation seems a way both to rule out symmetric 'explanations' (anyway, where these are unwelcome) and irrelevant 'explanations'" (p.728). Similarly, the "purported explaining facts introduce features which are causally irrelevant to what is being explained" (p.728). Timothy McCarthy, however, has pointed out some problems with causal conditions [McCarthy example, p.729]. Ruben suggests a solution inspired by Achinstein> "*The relevant premiss in McCarthy's arguments would say, for example, not only that c occurs, but also that c is the cause of e*" (p.731).

Railton (p.746)

Railton, in this paper from 1978, considers the world as a machine: "a vast arrangement of nomic connections" (p.748), which is not deterministic> "Some things *do* happen by chance, according to the dominant interpretation of our present physical theory, the probabilistic interpretation of quantum mechanics. Nonetheless, they can be explained: by that theory, in virtually the same way as deterministic phenomena - deductive-nomologically ... indications are that physical indeterminism is irremediable, and that the universe exhibits not only chances, but lawful chances" (p.746). "Yet it will not do simply to add to the D-N model a requirement that the explanans contain causes whenever the explanandum is a particular fact. First, some particular facts may be explained non-causally, for example, by subsumption under structural laws such as the Pauli exclusion principle [No two fermions can be in the same quantum state.]. Second, even where causal explanation is called for, the existence of general, causal laws that cover the explanandum has not always been sufficient for explanation: the search for explanation has also taken the form of a search for mechanisms that underlie these laws" (p.747). Moreover, "...explanations must be more than potentially-predictive inferences of law-invoking recipes. Is the deductive-nomological model of explanation therefore unacceptable? - No, just incomplete" (p.748).

Railton has problems accepting Hempel's inductive-statistical (I-S) arguments and their relativizations> "What I take to be the two most bothersome features of I-S arguments as models for statistical explanation - the requirement of high probability and the explicit relativization to our present epistemic situation ... derive from the inductive character of such inferences, not from the nature of statistical explanation itself" (p.751). Railton joins Jeffrey in his critique of Hempel's use of I-S arguments, as exemplified by> "...why should it be explicable that a genuinely random wheel of fortune with 99 red stops and 1 black stop came to a halt on red, but inexplicable that it halted on black?" (p.752). But Railton also finds Jeffrey neglecting rare ('practically negligible') events in his treatment> "Virtually impossible events may occur, and these deserve and can receive the same explanation as the merely improbable or the virtually certain" [cf. catastrophe theory] (p.752). Railton's example is alpha-decay in radioactive elements, in presenting his 'deductive-nomological-probabilistic model' as an extension of the D-N model. This D-N-P model applies to 'genuinely indeterministic processes'> "What must be given up is the idea that *explanations* can be based on probabilities that have no role in bringing the world's explananda about, but serve only to describe deterministic phenomena" (p.761-2). "The issue of showing the explanandum to have high (relative or absolute) probability is a red herring, distracting attention from the real issue: the truth or falsity, and applicability, of the laws and facts adduced in explanatory accounts" (p.764).