LAWRENCE SKLAR

THE CONTENT OF SCIENCE, THE METHODOLOGY OF SCIENCE AND HEMPEL'S MODELS OF EXPLANATION AND CONFIRMATION

(Received in revised form 20 November 1998)

1.

Peter Hempel's generosity of spirit and sympathetic warmth are, of course, legendary. Let me add to the many examples others have related of how his life so deeply and positively affected their own by telling of my first meeting with him.

In the middle of my first semester as a graduate student in theoretical physics at Princeton the growing awareness that I was not cut out for the life of a researcher into relativistic quantum field theory flashed into brilliant illumination: I needed out – and quickly! I crossed the short distance from Palmer Laboratory to McCosh Hall where Peter's office was located and called on him to ask if there was any possibility whatever that I might be considered for transfer into the philosophy program – despite the facts that I had little formal training at all in philosophy and had made no formal application to the program. At best I expected some advice on how to start the procedures rolling so that I could apply for such a change of program, and, perhaps, some encouragement that there might be some possibility that I would be successful in my application. Needless to say it was with utter astonishment that I listened to his response to my sad tale of failure as a physicist and desire to change my field of studies: Of course, he said, I was welcome to the philosophy program. But didn't I want to sit in on a few seminars first to make sure that this was a change I really wanted to make!

The warmth of this initial welcome into the philosophy program was matched by the continuing support and encouragement that

Philosophical Studies **94:** 21–34, 1999.

followed. Peter's deep and abiding concern for the welfare and the development of his students was always accompanied by his profound intellectual openness of mind and his genuine enthusiasm for the work of those around him, even when that work went off in directions very different from those marked out by his own so important work in scientific method. I have been fortunate to have come to know many in the profession of philosophy over the years who have combined deep philosophical research with fine teaching abilities and with true decency as human beings. But I have never met another Hempel.

2.

Two of Hempel's justifiably renowned contributions to the methodological philosophy of science are his detailed explication of explanation by subsumption of correlated events under generalities, the deductive-nomological model expounded in his famous paper jointly authored with Paul Oppenheim, and his important exploration of the notion of confirmation by instantiating instance.

Both contributions are replete with important insights. The careful outline of the notion of explaining events by showing them to be expectable conditional on the occurrence of other events given empirically established lawlike regularities placed this Humean idea in a modern context. It carefully delineated the formal and empirical requirements for an argument form to constitute a legitimate such subsumption having explanatory weight. It emphasized the place of explanation in a general framework of the scientific search for prediction and control. And it uncovered a number of hidden puzzles latent in the intuitive notions. By laying out this empiricist account of explanation with such care and precision, Hempel also provided the challenge against which all of those who found explanation as subsumption under generality not explanation enough had to contend.

The analysis of instantial confirmation once again took a highly intuitive idea, that hypotheses are confirmed by their positive particular instances, but carried the analysis of that idea to new levels of sophistication and insight. Starting from Nicod's naive idea of a generalized simple conditional being confirmed when protasis and hypotasis each were positively instanced by the same element (All A's are B's confirmed by something that is both an A and a B), Hempel showed how one could seek for a notion of confirmation by positive instance that was far more generally applicable. The most striking feature of the development was Hempel's brilliant discussion of the intuitive criteria of adequacy that one would want the ultimate formal definition to satisfy. Showing how a consequence condition and a converse consequence condition together led to a degeneration of the notion of indirect confirmation into triviality opened a realm of puzzles for those enamored of a hypothetico-deductive model of the confirmation of theories. And showing how as apparently innocuous a condition as that logically equivalent hypotheses ought to be equally confirmed by one and the same piece of instantial data led to the notorious "Paradox of the Ravens" and opened up new problem areas for research all on its own. In particular Hempel's analysis forced careful attention on such issues as the relationship of a purely qualitative notion of instantial confirmation to quantitative notions of confirmation, and it forced attention on the issue of the many factors to which confirmation of hypothesis by data may be argued to be relativised (such as background knowledge and the set of alternative hypotheses under consideration).

3.

One noteworthy feature of both the model of explanation and the model of confirmation Hempel constructed is their abstractness. They are constructed using a very sparse set of concepts and they are to be applicable in any context whatever. Nothing in the way of our scientifically developed understanding of the actual fundamental features of the physical world is invoked in describing the structure of what it is to be an example of an explanation or an example of a confirmation. And no restraints are placed on the domain in which the concepts can be applied. All explanations, in physical theories, in psychology, in the social sciences, are to be taken as subsumptions of correlated events under generalities, and nothing more is required of an argument form for it to constitute an explanation than that it has the appropriate logical form and that the explaining assertions are properly non-trivial. And all confirmations, in any domain of experience, are to be thought of as either direct, in having the predicates of an hypothesis instantiated by particulars in the right way in some evidential situation, or as indirect, mediated by the logical relation of hypothesis to some already confirmed hypotheses.

There are, I suspect, many reasons why Hempel strove for such abstractness in his accounts of both explanation and confirmation. Let us focus on the model of explanation and seek some of these out.

There is, of course, the direct ancestry of the deductive-nomological model in Hume's account of causation. To explain is to give causes, or at least so the discourse went at least as far back as Aristotle. But if causes are nothing but events spatially and temporally contiguous to the event caused and linked to it in the world by constant conjunction and in the mind by the imagination which induces the leap from the idea of the cause to that of the effect, then when we ask for the explanation of an event in the world, surely all we can be asking for is the linkage between that event and others to which it is related by the fact that the mutual occurrence of the events forms a general regularity in the world.

Now, of course, one thing Hume was out to do was to get any idea of "necessary connection" between cause and effect out of the world and into, at best, psychological propensities to believe. But, I believe, even by the time of Hume there was something else going on that was tending to push the philosopher into offering an account of causation and explanation that was so highly abstract. And this requires looking not just at the history of philosophy but at the history of science as well.

Already by Hume's time physics had suffered at least two "methodological shocks" concerning explanation. By this I mean that transformations of the theoretical understanding of what the world was like had resulted in profound transformations in the very idea of what it would be to offer an explanatory account of the nature of the world. The notion of what sort of thing science ought to be providing us when it provides us with explanations could be overthrown when a major theory that provided explanations of a kind earlier accepted as being the right sort of thing, the sort of thing explanations ought to be, was superseded.

With the rise of Copernicanism came the need for a new dynamical theory compatible with its postulation of a rapidly spinning earth. Galileo, building on the impetus theorists, first offers a kind of inertia taking natural, unforced motion of things on the earth to be the sharing of its circular motion. Descartes realizes that it is, rather, motion at constant speed in any straight line that is natural and unforced. But Descartes and his fellow new physicists invoke, along with their anti-Aristotelian theory of inertia, an altogether anti-Aristotelian theory of what it is to explain the phenomena of motion. Aristotelian notions of potentialities and powers, of natural places in the universe, and of final causes are all to be discarded on the rubbish heap of outworn metaphysics. Instead all is mass and extension. All motion is relative motion of one body with respect to some other, and all causation is causation by contact and displacement. A theory of the motion of the heavens, with the heavenly bodies carried along by the fluid plenum in which they are all immersed, is accompanied by a general rejection as non-explanatory of anything that fails to explain by positing motion of one body transmitted directly to a body with which it is in contact.

The brevity of the sway of the Cartesian model of explanation was truly astonishing. It was only a few years before Newton had demolished the entire Cartesian astronomical edifice. In the Newtonian scheme, the heavenly bodies move in empty space, and the forces that move them are due to almost Aristotelian powers of bodies that can act at a distance. Motion is not merely relative motion of one body with respect to another, but, instead, absolute motion with respect to "place itself". Time also is absolute, in that the accuracy of clocks is to be judged not relative to one another, but to the flow of time itself.

There are very curious aspects to the new scheme proposed by Newton to replace the Cartesian picture. Although Newton is chastised by the Cartesians for what positivists might call "metaphysical" elements – his postulation of absolute space and his tolerance of intrinsic causal powers – he is, at the same time, a proponent of the kind of thinned down explanation as subsumption under correlation that appeals to the empiricist temperament. For in response to the Cartesian accusation that he has failed to explain the action of gravity he replies, famously, “Hypotheses non fingo”, and comes close to maintaining that to give the correct law under which the observable phenomena fall is to do explaining enough. There is also a positivist tinge to Newton’s views on confirmation. Well aware of the dangers of “inference to the best explanation”, or the hypothetico-deductive model, he claims, dubiously to be sure, to have arrived at his generalities by induction from the phenomena alone. (Thereby, perhaps, serving as one of the inspiration for Hume’s skeptical attack on induction as no better off than the method of hypothesis at assuring reasonable belief.)

What is important to note here is that the Cartesians are not just proposing a physical theory incompatible with Aristotelianism. And Newton is not just proposing a physical theory incompatible with Cartesianism. In both cases the very idea of what it is for a theory to offer an explanation of the phenomena is up for question.

A second scientific debate carrying with it a debate over the proper methodology of science and the proper notion of what it is to be an explanation took place in the nineteenth century. This controversy, and its putative resolution, was probably the most direct historical influence on the deductive-nomological model. The conflict was the debate between “mechanicians” and “energeticists”. It is a curiosity that in this debate Newtonian ways of explaining things were often referred to as “mechanical”, given that in the earlier dispute with the Cartesians it was Newton who opposed the then-styled “mechanistic” account of the world in terms of transmission of motion by contact alone.

In the nineteenth century the two disciplines of the theory of heat, thermodynamics, and the theory of electricity and magnetism, electrodynamics, came into their fruition. Why did the laws of heat hold and why did the laws of electromagnetism hold? For the energeticist asking such questions could amount to nothing more than finding the most general and simple principles from which the laws of the respective disciplines could be systematically derived. The First and Second Laws of Thermodynamics and the unified Maxwellian laws of electromagnetism were all that one could ask of an explanatory theory of heat or of electromagnetism.

For the disciple of mechanism, however, one could ask much more. One could ask for the underlying mechanisms that "explained" why the phenomenological laws held. In the case of the theory of heat, this led to kinetic theory and ultimately to statistical mechanics, in which the macroscopic thermal behavior of things was accounted for in terms of the mechanical interactions of its microscopic components supplemented by important probabilistic posits over initial conditions. For electromagnetism it led to the attempts at finding mechanical models of the aether, in order to construct a mechanism that could support electromagnetic waves with their well known fundamental features (such as polarization that implied strict transversality to the wave magnitude). The mechanical theory of heat, of course, became our orthodox theory, whereas the mechanical theory of electromagnetism remains as nothing more than an historical curiosity.

But it was, I think, the failure of mechanism in the case of electromagnetism, rather than the success of mechanism in the theory of heat, that influenced methodology the most. Despite the overwhelming success of the atomic theory of matter, and the hypothesizing of deeper and deeper "hidden" levels of being to explain what happens at the level of the observable, it was the positivism of Mach and Duhem that emerged victorious in the battle of the methodologists. The orthodox doctrine became that which eschewed "hypotheses", meaning the real postulation of domains of unobservable things and properties as explanatory of the observable, except insofar as the hypotheses could be reconstrued as convenient "economical" summaries of the regularities among the observables. And the orthodox doctrine of explanation became the neo-Humean account in which to explain could mean only to subsume the welter of phenomena under generalizations that expressed constant correlations.

The refusal to countenance notions of "cause", despite labeling the particular conditions of the explanans by "c's", was motivated not simply by Humean skeptical attacks on alleged early notions of causation, but by claims to the effect that in the new physics all that ever appeared were lawlike generalizations associating phenomenon to phenomenon, and that the notion of the cause of a happening, carrying with it as it did much outworn metaphysics and many inapplicable "common sense" ways of accounting for the world, never appeared in any role in this new physics at all.

4.

No sooner had the deductive-nomological model appeared than the attacks on its adequacy began. Some of these argued that the model demanded too much of an explanation. Hempel and Oppenheim had, of course, always acknowledged that a class of probabilistic or statistical explanations existed that could not fit into the deductive-nomological model. Other, however, challenged the need for generality of any kind in an explanation at all. Where, they asked, were the generalizations in historical explanations? Some argued for the appearance of generalization like components in some explanans, but generalizations that were not, as those demanded by the deductive-nomological model had to be, empirically supported. There were, it was claimed, rational action explanations where the invoked generalizations were grounded not in induction or empirical generalization but in rational means-ends calculations. Then there were those who pointed out the pragmatically determined contexts in which explanations were demanded and supplied, and argued that these required that relevance considerations be added to the other demands of the deductive-nomological model.

Possibly the most persuasive objections to the deductive-nomological model as a full characterization of explanation in science, though, are those that criticize the model for allegedly leaving out of the picture essential components required of an scientific explanation. The criticism starts with the familiar counter-examples to the analysis, propositions arranged in the models premise-conclusion form and satisfying its other constraints such as essentially invoking at least one empirical generalization, but where we instinctively feel that such an argument form hardly counts as an explanation. There is the problem of the model apparently alleging that the length of the shadow explains the length of the flagpole. There are surface generalizations that seem to express "mere" correlations that hardly seem explanatory. And there are the cases where explanans and explanandum event are in the "wrong" time order.

What is the deductive-nomological model missing out on when it declares such argument forms explanations but we don't find them such? The natural response is that the model fails to realize the importance of "causality" and "mechanism" in explanations. Finding mere correlations, subsuming events under generalities no matter how bold and how well empirically established, isn't finding explanations, it is sometimes said. To explain an event is to reveal its causes. To explain a process is to unveil the hidden mechanisms underlying its regular behavior. Such debates about the necessary components of an explanation outrun the abstract concerns of the philosopher, of course. Witness the continual courtroom battles as lawyers deny that epidemiology is enough to establish causality and that it is causality that the law requires for responsibility. Without causation, without mechanism, they say, we have no explanation for what has happened, and it is explanation that the law demands.

5.

But how are we to deal with the repeated claims that models of explanation that invoke mere subsumption of events under generality leave out the crucial causal and mechanistic aspects of explanation? One thing we can do is to look for some understanding of the notion of causation that is almost as abstracted from the specific contents of our actual scientific beliefs as is the deductive-nomological model. We can, for example, explore our intuitions about how causal judgments on our part are intimately related to our judgments about what would have happened or not happened had some condition in the world been other than it was. That is we can pursue the relationship of causation to counter-factuals. And we can go on from there invoking possible worlds, their relations of proximity and similarity, and all the familiar apparatus of the metaphysics of causal modalities.

But when we have done all that will we still be, in the end, mystified by what our intuitions about causation and mechanism are grounded on other than our understanding that, in the abstract, there is regularity to the world and the gaining of knowledge of this regularity is essential for our gaining of predictive and controlling powers. That is, what is at base pushing us to insist that there is more to causation and mechanism than mere regular order?

The abstract nature of the deductive-nomological model is, as noted, in part a result of the historical drive to eliminate any preconception dependent upon the maintaining of particular scientific views from our model of the method of science itself. Too often we found ourselves wedded to a model of what it is to do proper science that was constricted by the specific nature of the science of the time. Consider the Cartesian reluctance to accept Newton's dynamics and theory of gravity as really explanatory of the motion of the planets, or the nineteenth century atomist-mechanist's refusal to allow that thermodynamics and electromagnetic theory might be in no need of a mechanical underpinning in order to explain our world to us. But can we really so neatly distinguish the "pure" methods of science from the specifics of what is contained in our accepted fundamental physical theories? Can we eliminate from our models of methodology the specific nature of the world present to us pre-theoretically that these fundamental scientific theories describe?

The constant demand for "showing the mechanism" in explanations can, perhaps, be understood in terms of intuitions on our part about what an explanation needs to be, intuitions that are deeply rooted in either our familiar experience of what kinds of regularities we ordinarily encounter, or in what kinds of specific fundamental theories seem so deeply embedded in our science that we can hardly imagine them eliminated in any future scientific change. As an example of the former, consider the Cartesian demand that all explanation of change of motion must be grounded in contact force and displacement of one object by another. Surely this is an intuition about explanation that comes from the pervasiveness of our everyday experience about how things do get put into motion in our usual world (assuming, that is, that we don't play with magnets all that often!). As an example of the latter, consider the nineteenth century mechanicians, living in a world of science that for two-hundred years had been so dominated by Newtonian notions of masses and forces acting upon them and of the interactions of matter in motion.

Perhaps the clearest context in which the role played by the specific constitution of nature as we encounter it in experience, and the role played by our theoretical understanding of the origin of that specific aspect of the way nature is, in the generation of our very understanding of what it is to give an explanation of a phenomenon, can be found in exploring the origin of our idea that in causal explanations that which explains precedes in time that which is to be explained. The simplest objections to the deductive-nomological model that allege its failure to take causation into account are those constructed by looking at two events that are lawlike conjoined and arguing that while one explains the other, the explanation in the reverse order is blocked by the asymmetry in temporal order of the events. Past explains present and future, future cannot explain present or past, but the deductive-nomological model allows for no such asymmetry (even though that asymmetry was just plugged in by its Humean ancestor).

We don't really know where our idea of a time asymmetry in explanation comes from. Perhaps it is epistemic in origin, related to the fact that records and memories exist of past but not of future. But, perhaps, that asymmetry as well as the explanatory asymmetry have their origin in some actual asymmetry in time encountered in experience and accounted for in fundamental physics. Perhaps, the argument often goes, it is the asymmetries in time described by thermodynamics and other related "dissipations of order into disorder into the future" that underlie the intuition that explanations go from past to future.

Here we need not explore the differing details of accounts that move in this direction, say those of Reichenbach resting on principles of common cause or those of Lewis that seek for asymmetries in our evaluations of the truth of forward- and back-tracking counterfactuals. What matters here is that we have claims that, were they true, would cast some doubt on the very idea that a methodological program, such as analyzing the nature of explanation in science, can be carried out in total abstraction from, and studied indifference to, the specific actual general aspects of the world in which we live and the specific structures of the theoretical accounts we can construct to deal with that world. If such a fundamental intuition of ours about explanations, that is that they should be causal and that being causal implies explaining later happenings by reference to earlier ones, rests upon "contingent" features of the world of our everyday experience that are to be accounted for in terms of the specifics of our foundational physical theories, then there is good reason to think of methodology as inextricably entangled with the ongoing program of trying to understand the most fundamental structural features of our world within science itself.

All of this being said, there does seem to be something special about the core idea of the deductive-nomological model, that is that explanation is subsumption under generality. One wonders if there is some "contingent" feature of the world that itself grounds our very most basic idea that to explain is to subsume under regularities. What could a world be like in which our notion of explanation didn't have that as at least a fundamental component? Would there be any place in such a world for anything like what we call explanation at all?

6.

An examination of Hempel's notion of instantial confirmation may lead to similar observations about the inextricability of methodology and empirically contentful foundational science. Here at least one route to this conclusion comes from the repeated realization, at least since Whewell, that we have difficulty in making sense of any confirmational notions without presupposing a privileged group of substantive and attributive properties as characterizing things and events in the world.

Over and over again we have been taught that it is hard to give any formalization of some notion of confirmation that doesn't take as a given a specified set of designated classificatory categories. Arguments to this effect can be seen in Whewell's critique of Mill's methods, in Bertrand's "paradoxes" for attributing probability and their later incarnation in the "linguistic relativity" of Carnap's logical probability, in Goodman's justly famous critique of "the future will be like the past", and, perhaps, in a transmuted form, even in Wittgenstein on rule following.

The bulk of the critical discussion of Hempel's qualitative instantial confirmation did not focus on such issues, but, rather, on discussions of the criteria of adequacy, in particular on how to reconcile hypothetico-deductive reasoning with the obviously disastrous consequences of adopting a general converse consequence condition, and on finding a way out of the "Paradox of the Ravens". But the problem of "appropriate kinds" is latent in Hempel's analysis as well.

Where do our preferred predicates come from? Goodman suggests entrenchment, but other realists suggest: From our best available scientific theories. It is these theories themselves that tell us when it is, and when it isn't, reasonable to apply our inductive methods, whatever they may be, to things classified in one way or another. If this is right, then we may have here another demonstration of the inextricability of methodology from the task of finding within science itself the appropriate framework for describing the world.

7.

The desire on the part of logical empiricists and logical positivists to frame methodology solely in terms of abstract logic, and to characterize it as above the fray of the specifics of current science – even foundational science – had its deep motivation. All too often they had seen methodology limited by the presupposition that the science of the day limited the range of any possible science. If Aristotelian, Cartesian or Newtonian science made claims to universality as a foundational science, it was all too tempting to think that the way that accepted foundational science explained constituted what explanation had to be in general.

But the continual revolutions of science, with new disciplines being created that bore little conceptual overlap with the old, and with the repeated overthrow of the old even in its preferred domains, put great pressure on methodologists to frame notions of explanation and confirmation that would be independent of the presuppositions of the specific science of the day, even the specific most general and foundational science of the day.

All of this lies, I think, in the background of such abstract characterizations of explanation as the deductive-nomological model and the instantial notion of qualitative confirmation developed by Hempel. I don't think we yet really understand the place of such abstract methodology in our overall picture of the world. Can method be so disassociated from the specific content of the theories that, allegedly, are framed within the method in question? Or do method and content form an inextricable whole?

Department of Philosophy
University of Michigan
2215 Angell Hall
435 S. State St.
Ann Arbor, MI 48109-1003
USA