

The Causal Mechanical Model of Explanation

Wesley Salmon's *Scientific Explanation and the Causal Structure of the World* (*SE*) presents a sustained and detailed argument for the causal/mechanical conception of scientific explanation which Salmon has developed in a series of papers over the past decade. *SE* ranges over a wide variety of topics – in addition to the material discussed below, it contains discussions of probabilistic theories of causality, scientific realism, of the notion of objective reference class homogeneity, and of much of the recent literature on causality and explanation. In my judgment, *SE* is the most interesting general treatment of scientific explanation since Hempel's *Aspects of Scientific Explanation*. Even those who are not persuaded by many of Salmon's conclusions will find this book eminently worth careful study.

According to Salmon, causality plays a central and fundamental role in scientific explanation. To explain a particular occurrence is to show how it “fits into” the causal network of the world (276). Causality has three fundamental “aspects” (179). The most basic causal notion is that of a *causal process*. A causal process is characterized by the ability to transmit a mark or the ability to transmit its own structure, in a spatio-temporally continuous way.¹ Examples include the movement of a free particle or an electromagnetic wave through empty space. Causal processes are to be distinguished from pseudo-processes (e.g., the successive positions occupied by a shadow cast by a moving object) which lack the above abilities. *Causal interactions* occur when one causal process (spatio-temporally) intersects another and produces a modification in its structure. A typical example is a collision of two particles. Interactions commonly involve correlated changes in structure governed by conservation laws. *Conjunctive forks* involve correlations among spatio-temporally separated effects which are explained in terms of separate causal processes deriving from a common cause, as when food poisoning is invoked to explain the indigestion experienced by a number of people who attended a picnic.

I would like to thank Paul Humphreys for helpful comments on an earlier draft of this paper.

A scientific explanation of some particular outcome will consist of citing (some portion of) the causal processes and interactions leading up to that outcome. The explanation of a type of outcome or a generalization is a derivative notion; such an explanation will describe what the particular causal processes and interactions responsible for instances of that type or generalization have in common. Purported explanations which fail to cite genuine causal processes and interactions – Salmon suggests thermodynamic explanations which appeal to the ideal gas laws as cases in point – are spurious or at least seriously defective precisely because of their noncausal character. Salmon claims that it is a virtue of his approach that it enables us to see why such noncausal explanations are defective, while the traditional DN model does not.

While Salmon thus no longer subscribes to the idea, expressed in his well-known monograph “Statistical Explanation and Statistical Relevance,” that information about statistical relevance relations is by itself explanatory, his discussion in *SE* makes it clear that he by no means wishes to entirely abandon the SR model. Instead Salmon suggests that scientific explanation is typically a “two-stage affair” (22). In the first stage, one assembles information about statistical relations of the sort demanded by the SR model; in the second stage, one provides a causal account in terms of processes and interactions of why those relations obtain. Information from the first stage is said to play the role of providing evidence for claims about causal connections introduced in the second stage, but it is only claims of the latter sort which have explanatory import. Providing a correct characterization of the SR basis, as Salmon now prefers to call it, is accordingly still an important part of the project of providing an adequate model of scientific explanation, and Salmon devotes a substantial part of *SE* to developing such a characterization and defending it against objections.

While this new formulation improves in interesting respects on Salmon’s earlier formulation, the fundamental idea of the earlier account is retained: one begins with a reference class A and a set of explanandum properties $\{B_i\}$ and then, by introducing a set of statistically relevant factors $C_1 \dots C_s$, partitions the reference class A into a set of mutually exclusive and jointly exhaustive cells $A.C_i$; each of which is required to be “objectively homogeneous” with respect to the set $\{B_i\}$ – i.e., none of the cells “can be further subdivided in any manner relevant to the occurrence of any B_i .”² (37) Moreover, Salmon still wishes to retain a number of characteristic features of the SR model – for example, he still thinks of statistical theories like elementary quantum mechanics (QM) as providing explanations of individual outcomes in circumstances in which they merely permit a calculation of the probability with which those outcomes will occur. Thus, according to Salmon one can explain why some individual carbon-14 atom decayed during a certain short time interval by citing, among other things, the probability of decay during this interval, even though this probability will be quite low. Various alternative theories of explanation (such as Hempel’s IS model) are still

judged as inadequate in part because of their inability to accommodate supposed explanations of this kind.

Salmon's discussion takes place against the background of a provocative contrast between different general conceptions of scientific explanation. *Epistemic conceptions* represent attempts to characterize explanation in epistemic terms—the best known example is the DN model, in which the key epistemic notion is nomic expectability. While the epistemic conception, at least in the DN version, leads one to think of an explanation as an argument or an inference, and accordingly to look for a “logic” of explanation valid for all possible worlds, Salmon holds that this is a “futile venture, and . . . little of significance can be said about scientific explanation in purely syntactic or semantic terms” (240). By contrast, the *ontic conception* which Salmon favors represents an attempt to characterize explanation in terms of the fundamental causal mechanisms which, as a matter of contingent fact, operate in our world, with no suggestion that similar explanatory principles would be useful in all logically possible worlds.

In what follows I shall explore four general questions raised by Salmon's discussion: (1) Is it reasonable to suppose that all scientific explanations will meet the requirements of the causal/mechanical model? (2) Does Salmon's model capture all of the features which are relevant to the assessment of scientific explanations? (3) Should we take statistical theories like quantum mechanics as providing explanations of individual outcomes; and, relatedly, does Salmon's SR basis capture the way in which statistical evidence is relevant to the construction of explanations in QM and elsewhere in science? (4) Should we abandon epistemic conceptions of explanation in favor of Salmon's ontic conception? Is there a single logic of explanation valid in all possible worlds or, for that matter, everywhere in our world?

I

Salmon's conception of explanation seems to fit most neatly simple physical systems, whose behavior is governed by the principles of classical mechanics and electromagnetism (including such common-sense paradigms of causal interactions as the collision of a golf ball with a tree limb). I think it is clear that Salmon has captured a number of the central features of the notion of causation, as it is applied to such systems. In connection with such systems one typically thinks of causation as involving the transfer of energy and momentum in accordance with a conservation law, and it is a consequence of Special Relativity that if, when transferred, such quantities are conserved, they must be conserved locally—that there is no causal action at a distance and that causal processes will be spatio-temporally continuous, just as Salmon claims.³

However, even in physical contexts, the idea that explanation must involve the tracing of causal processes and interactions in Salmon's sense becomes increas-

ingly problematic, as we move away from the above paradigms. Consider, for example, explanations which make reference to the geometrical structure of spacetime as in Special and General Relativity. Even if one thinks, as Salmon does, of a particle moving along a time-like geodesic as a causal process, it seems fairly clear that the *explanation* one gives in General Relativity for why the particle moves as it does is not a causal explanation. The explanation one gives will make reference to facts about the affine and metrical structure of spacetime and the variational equation of motion $\delta J_d = 0$, and these facts about the geometrical structure of spacetime will in turn be explained in terms of the distribution of mass and energy (as expressed in the stress-energy tensor). Neither the fact that the particle moves along a geodesical path nor facts about the geometrical structure of spacetime are themselves explained in terms of continuous causal processes or interactions between such processes, or in terms of such characteristically causal notions as forces, or transfers of energy and momentum. A similar point holds with respect to the explanations of length contraction and time dilation in a moving inertial frame provided by Special Relativity. A rapidly moving clock is a continuous causal process, but the time dilation and length contraction it will exhibit have their explanation in the structure of Minkowski spacetime and are geometrical effects rather than causal effects whose origin is to be sought in causal processes and their interactions. Indeed the contrast between the explanation of such phenomena in Special Relativity and their explanation in a theory like Lorentz's, in which these effects are regarded as a consequence of the operation of electromagnetic forces, seems to be precisely the contrast between a non-causal, geometrical explanation and a causal one.⁴

Explanation in elementary quantum mechanics represents a second class of explanations which do not fit comfortably within Salmon's framework. Salmon is aware of this fact and includes in *SE* an interesting discussion of explanation in QM. This focuses largely on the difficulties EPR type correlations create for the common-cause explanatory principle. Salmon seems to suggest that it is an explanatory deficiency in QM that it fails to explain such correlations in terms of spatio-temporally continuous processes connected to a common cause and that it is an advantage of his account of explanation that it allows us to recognize this, while the standard DN model would not.⁵

This is hardly the place to embark on a detailed treatment of the notoriously difficult and intractable issues surrounding the interpretation of quantum mechanics. I confine myself to two general observations concerning Salmon's discussion. The first is that it is at least arguable that any plausible theory of scientific explanation must make room for a possibility which Salmon seems to reject: that we can learn, as a result of empirical inquiry, not just that various particular candidates for an explanation of some phenomena are defective, but also that the demand for any explanation (or at least any explanation of some very general type) is inappropriate or misguided. In many cases the question of whether the demand

for a certain kind of explanation of some explanandum is a reasonable one is not something we should expect to settle just via appeal to an abstract model of explanation (whether of the deductive-nomological or causal/mechanical variety) but will depend rather on the results of empirical inquiry. There is thus a difference between a theory that fails to successfully explain some potential explanandum, where it is agreed that the demand for such an explanation is perfectly appropriate, and a theory that provides us with good reasons for rejecting the demand for such explanations—only in the former case is talk of an explanatory lacuna appropriate.⁶

It seems to me that one of the central issues raised by Salmon's discussion is whether the former or the latter characterization is the more appropriate way to think of the failure of quantum mechanics to explain phenomena like EPR correlations in terms of a common-cause mechanism. This issue is in turn bound up with more general issues concerning the interpretation of quantum mechanics. Although Salmon does not explicitly commit himself to a general view about the interpretation of QM, much of his discussion seems to suggest or presuppose a rejection of the standard, Copenhagen interpretation of the theory in favor of a "realistic" interpretation according to which all measurable quantities have precise values prior to or independently of measurement.⁷ Given the expectations generated by this sort of realism, it is particularly natural, if not irresistible to inquire, as Salmon does, after the details of the mechanism by which nature arranges things so that the EPR correlations hold and to suppose that it is a limitation on the explanatory power of quantum mechanics that it fails to specify such a mechanism.

However, as is well known, various no hidden variable theorems seem to show in principle that no common-cause mechanism and indeed no realistic theory of the sort that Salmon envisions could exist which reproduces the experimentally observed correlations.⁸ It is tempting to think—I take this to be the attitude of many physicists—of such impossibility results as one of a number of considerations which show that there is something misguided or inappropriate about the demand for a common-cause mechanism of the sort described above. Rather than concluding that there is an explanatory gap in quantum mechanics, those who adopt this view will conclude instead that it is our theory of explanation and our prior expectations that requests for explanation of a certain general kind are always in order which need revising. More needs to be said by Salmon to support his alternative characterization of the failure of QM to explain EPR type correlations and to show that this characterization can be motivated without the assumption of a kind of realism which seems empirically false and which would be rejected by the majority of physicists.

The second and more general observation is that because Salmon's discussion focuses largely on the EPR problem, the implications of his views for the assessment of the explanatory power of QM generally are not drawn as sharply as they

might be. Consider an elementary textbook example, referred to briefly by Salmon: the derivation of the probability that a charged particle will penetrate a potential barrier by solving the time-independent Schrödinger equation for the system. This strikes me as a paradigm of quantum mechanical explanation—it represents the (highly idealized) basis for the usual elementary treatment of Salmon's often repeated example of the radioactive decay of a carbon-14 atom. However, as Salmon acknowledges (256–57), this explanation certainly does not seem to involve the explicit tracing of continuous causal processes leading up to the penetration of the barrier. If we adopt Salmon's model of explanation, it is not clear that we can avoid the conclusion that QM is nonexplanatory in this sort of typical application (and indeed in virtually all of its applications), and not just that it fails to explain EPR type correlations. Various scattered remarks suggest that Salmon is perhaps willing to accept this conclusion, in part because he thinks (or at least holds that it is not obviously false that) more recent theoretical developments in quantum electrodynamics (QED) and quantum chromodynamics (QCD) more closely approximate the demands of his model.

This strikes me as unpersuasive for several reasons. First, I think that most physicists and chemists think of elementary, nonrelativistic QM as a satisfactory explanatory theory as it stands in the domains to which it is taken to apply and that it does not require supplementation by some more causal theory (in Salmon's sense) before it becomes explanatory. Quite apart from this, it is unclear how a theory like QED, which employs such notions as virtual particles and vacuum fluctuations or in which one does calculations by summing over all possible past histories of a particle or over all possible ways in which a certain interaction can occur, is one in which, as Salmon claims, a "broadly causal picture seems to emerge" (255). Salmon's brief remarks in support of this claim (which he seems to intend just as tentative suggestions) are, I think, simply not detailed enough to enable one to see very clearly what he has in mind. Moreover, although the issues involved here are murky, the no hidden variable results alluded to above seem not merely to undercut the possibility of a realistic interpretation of quantum mechanics as presently formulated, but to strongly suggest that, however our understanding of microphysics may change, we should not expect a future theory in which phenomena like EPR type correlations are explained in terms of spatio-temporally continuous processes. Salmon seems (254–55) to envision a theory which provides causal explanations (in his sense) of such phenomena but which is at the same time not a local hidden variable theory, but it is not obvious how this is possible.

More also needs to be said about how Salmon's model applies to complex physical systems which involve large numbers of interactions among many distinct fundamental causal processes. In such cases it is often hopeless to try to understand the behavior of the whole system by tracing each individual process. Instead one needs to find a way of representing what the system does on the whole or on

the average, which abstracts from such specific causal detail. It is not clear how Salmon's model which is formulated in terms of individual causal processes and their binary interactions, applies to such cases.

Consider again Salmon's example of the explanation of the behavior of a gas in terms of the statistical mechanics of its component molecules and the contrast Salmon wishes to draw between this and a (noncausal and hence defective) explanation which appeals to the ideal gas laws. Plainly it is impossible to try to trace the trajectories and interactions of each individual molecule (the individual causal processes involved here) and to exhibit the behavior of the gas as in any literal sense the sum of these. Moreover, although some of Salmon's remarks perhaps suggest that he takes a contrary view, it seems equally clear that it is a trivial, nonserious explanation of the behavior of the gas to say simply that it is composed of molecules, that these collide with one another in accordance with the laws of Newtonian mechanics and that (somehow) the behavior of the gas results from this. The usual sort of treatment found in textbooks on statistical mechanics does neither of these things, but instead proceeds by making certain very general assumptions about the distribution of molecular velocities, the nature of the forces involved in molecular collisions and so forth, and then deriving and solving the Maxwell-Boltzmann transport equation for the system in question. One then shows how various facts about the behavior of the gas (such as that it obeys the ideal gas laws) follow from the solution of this equation.

This sort of treatment does not literally describe in detail the collisions of any individual molecule, let alone the entire collection of molecules comprising the gas. Instead, in constructing an explanation one abstracts radically from details of such individual causal processes and focuses on finding a way of representing the aggregate behavior of the molecules. In this treatment, such characteristically "epistemic" or "inferential" concerns as finding techniques for actually solving the relevant equation governing this aggregate behavior and for avoiding computational intractabilities are of quite central importance. Rather than merely mirroring facts about causal interactions and processes the relevance of which for inclusion in the explanation is determined on other grounds, such epistemic considerations seem to have an independent role in determining why this sort of explanation takes the form it does. One omits causally relevant detail about individual interactions for the sake of such epistemic considerations as derivability and generality.

Examples such as this raise a number of natural questions in connection with Salmon's discussion. Just what does the causal/mechanical model require in the case of complex systems in which we cannot trace in detail individual causal processes? How, in terms of Salmon's fundamental notions of causal processes and interactions, does the statistical mechanical explanation sketched above amount to successfully specifying a mechanism? How much of the above explanation (how much detail and of what sort) does one have to provide before one has

satisfied the constraints of the causal/mechanical mode? To what extent can one capture or account for the salient features of the above explanation without appealing to epistemic or inferential notions? On what grounds can we say—as I think that we must—that the trivial explanation described above fails to provide a mechanism or to explain the behavior of the gas? A fully worked out version of the causal mechanical model needs to provide a principled answer to such questions.

II

Even if we agree that some or all good explanations involve the postulation of continuous causal processes, it is doubtful that all of the features of an explanation which are relevant to its assessment have to do with the extent to which it successfully postulates such processes. There are important explanatory desiderata which are not captured by an account which just focuses on the idea that to give an explanation is to cite a cause. One such desideratum is that a good explanation should diminish one's sense of arbitrariness or contingency regarding the explanandum. For example, an important feature of many explanations in statistical mechanics (and indeed of explanations elsewhere that successfully invoke equilibrium-type considerations, as in evolutionary biology)⁹ is that they proceed by showing that for a great many possible sets of initial conditions, an outcome like the actual outcome would have ensued. In this way, one's sense that the actual outcome was fortuitous or arbitrary is at least partly removed. This feature is missed in an account like Salmon's which just focuses on describing the actual causal history leading up to the explanandum-outcome and leaves out questions about whether other possible but not actual histories would have led to a similar outcome.

The idea that a successful explanation ought to diminish arbitrariness is also reflected in the common distaste of scientists for theories that contain many free parameters, the values of which are not determined by the theory itself but rather must, as it is commonly expressed, "be put in by hand"—introduced with no other rationale than that they are required by the data. The idea that a theory having this sort of feature is defective as an explanation because it is left arbitrary why the parameters in question should have just these values, rather than some other set of values, is very common, especially among physicists. Various theories of the weak and strong interactions (including QCD) are frequently criticized on just this score and one of the central attractions of recent "superstring" theories is that they seem to go a long way toward eliminating this sort of arbitrariness.¹⁰ Here again, it is at least not obvious how this sort of dimension of explanatory assessment is captured by Salmon's model, with its focus on individual causal processes and its de-emphasis on the role of inferential considerations in explanation.

Another consideration which is quite central to the assessment of explanations

but is insufficiently emphasized in Salmon's model has to do with the idea that a good explanation will provide a unified account of phenomena previously thought to be unconnected.¹¹ In my view, the idea that such unification is an important goal in scientific explanation provides a much more natural account of the salient features of, say, contemporary high energy physics than the idea that theory construction in this area is driven by the demand that microphysical phenomena be explainable in terms of continuous causal processes. It is certainly the former notion and not the latter which is generally emphasized by physicists in their own accounts of their activities.¹² The demand for such unification is evident in such recent achievements as the Weinberg-Salam electroweak theory, which unifies the electromagnetic and weak forces, and in more recent proposals for a unified treatment of the strong and electroweak forces. One can plausibly point to many central mathematical features of such theories, such as the imposition of various gauge symmetry requirements, as providing a concrete implementation of this demand for unification.

Salmon, after criticizing epistemic versions of accounts of explanation which assign a central role to theoretical unification, suggests that his own treatment can also take unification to be an important aspect of explanation since on his view, "unity lies in the pervasiveness of the underlying mechanisms upon which we depend for explanation" (276). This makes it sound as though unification is an incidental (but welcome) byproduct of the search for causal mechanisms. I believe that this gets matters backwards—in many areas of science, it is the demand for unification which is primary, and one determines what the relevant mechanisms are in the light of this demand. Spontaneous symmetry breaking or the Higgs mechanism are regarded as important mechanisms in high energy physics because of the role they play in current unification programs and not because they are independently required by the constraints of the causal/mechanical model.

The significance of this last point becomes even clearer when we consider the implications of the causal/mechanical model for explanatory theorizing in biology, psychology, and the social sciences. A great deal of theorizing in these disciplines proceeds on the assumption that systems which may differ significantly from the perspective of some fine-grained, microreductive causal theory may nonetheless exhibit interesting common patterns or regularities at a more macroscopic level of analysis and that one can construct explanatory theories by focusing on such patterns and regularities. Put crudely, the basic idea is that complex systems can exhibit different levels of organization and that, corresponding to these, different levels of explanation are appropriate. Thus systems that differ in underlying physical or chemical respects can nonetheless be treated as similar for the purposes of biology, psychology, or economics.¹³ In the absence of such an assumption, it is hard to see how serious explanatory theories in these disciplines are possible.

Thus, explanations in evolutionary biology of why quite different organisms

possess similar traits or behaviors will often make reference to the fact that these organisms face quite similar adaptive situations or selection pressures, despite the fact that the proximate mechanisms underlying these traits may be quite different in the case of different organisms. For example, one finds, in evolutionary theory quite general game-theoretical explanations of various behavioral strategies (regarding defense of territory, parental investment in offspring, and so forth) although the immediate causal antecedents of such behavior in different organisms will be quite different. Similarly, consider such general results in evolutionary theory as Fisher's fundamental theorem on natural selection (that the rate of increase in fitness in a population at a time equals the additive genetic variance in fitness at that time) or the standard explanation of how heterozygote superiority can lead to a stable polymorphic equilibrium. Here too we have quite general results which apply to a wide range of different organisms acted upon by natural selection despite important differences of causal detail. Or consider the use of quite general formal models in ecology—say, the Lotka-Volterra equations for prey/predator interactions. Here the attempt is to model and explain quite general features of such interactions, despite the fact that the causal details will differ greatly from population to population. Finally consider explanations of people's cognitive capabilities (e.g., chunking and recency effects in memory or tendencies to make certain inferential errors) in terms of the way they process and store information. It is a central claim of much theory in cognitive psychology that such accounts can provide genuine explanations even though they do not describe in detail the operation of neurophysiological or biochemical mechanisms and even though similar information-processing strategies may have interestingly different neurophysiological realizations in different subjects.¹⁴

None of these explanations seems to explain by tracing in detail continuous causal processes or underlying physical mechanisms. Rather, as the statistical mechanical case, they explain by abstracting from such detail and finding general patterns. An account of explanation which attaches a central role to theoretical unification and nonarbitrariness (and which recognizes a connection between explanation and inference and is sensitive to the ways in which computational intractabilities can interfere with attempts to explain) can make sense of and legitimate the above explanations. The demand for explanations that unify (and diminish arbitrariness, and so forth) and the demand for explanations that specify the details of fundamental physical mechanisms or that trace continuous causal processes can conflict, and when this happens, it is often the former demand that ought to prevail. By contrast, I think it is at least not obvious (even if one leans *very* heavily on the idea that the causal/mechanical model is an "ideal" which is only partially realizable in practice or to which successful actual theories represent an approximation—see *SE*, 263ff.) how Salmon can avoid the conclusion that many of the above theories are pretty dubious as explanations, in virtue of their apparent failure to specify continuous causal processes.

It would have been very useful to have had an explicit discussion of such examples in *SE*. Does Salmon think (appearances to the contrary) that the above examples satisfy the requirements of the causal mechanical model?¹⁵ If so, what sort of biological and psychological explanations does the model rule out? Is there some natural way, within the context of the causal/mechanical model (even in its ideal text version) of avoiding the (seemingly wrong-headed) conclusion that the best way to improve the above explanations would be to add more specific detail about proximate mechanisms and continuous causal processes (even if generality and other explanatory desiderata are lost) and that the ability of the above treatments to unify disparate phenomena has little to do with their explanatory power at least if this unity is not a result of a sameness of underlying mechanisms? Or is this perhaps a conclusion which Salmon would wish to endorse and defend?

III

Like Hempel (and indeed most other writers who have discussed the matter) Salmon retains the idea, associated with the original SR model, that statistical theories like QM, which in many circumstances merely specify probabilities strictly between 0 and 1 that certain outcomes will occur, nonetheless can be thought to explain those outcomes.¹⁶ Salmon is quite effective in showing that *if* this idea is accepted it will have very important implications for how one ought to think about explanation. For example, Salmon argues convincingly that given this conception of statistical explanation, it is arbitrary to suppose (as Hempel did in his original IS model) that an explanation which assigns a high probability to some outcome is for that reason better than an explanation which assigns a low probability to some outcome. He also shows that given the above conception it will be hard to avoid the conclusion that the same explanans can explain both the occurrence of some event *E* and the nonoccurrence of *E* (113)—a conclusion that is certainly inconsistent with a great many philosophical accounts of explanation.

However, like other writers on the subject, Salmon devotes comparatively little attention to arguing for the antecedent of the above conditional—i.e., that statistical theories like QM *do* provide us with explanations of individual outcomes in the sense intended. From his brief discussion, it appears that he is in part influenced by the idea that the only plausible alternative to his own view is that statistical theories explain facts about approximate relative frequencies in large numbers of outcomes (e.g., that roughly 1/10 of the atoms in some large collection decay in a certain interval or that approximately 375 out of 500 plants have red blossoms (216)). I agree with Salmon that this alternative is inadequate. First, one wants such statistical theories to be applicable to or able to explain facts about small populations or individual outcomes. Second, given a theory which predicts (P) that an outcome *E* will occur with a certain probability *k*, the claim that the approximate relative frequency of *E* in even a large population will be

“close” to k does not of course follow deductively from (P), but is, according to the law of large numbers, merely probable given P. Thus to claim that facts about relative frequencies are explained in such a case is not to avoid the notion of statistical explanation, but to embrace a particular version of it—in this case, something like the IS model in which an explanandum is explained by finding an explanans which confers a high probability on it. If this were the only alternative to Salmon’s approach, his treatment would be vindicated.

It seems to me, however, that Salmon fails to emphasize a rather natural third possibility: one can distinguish between claims about probabilities and claims about relative frequencies (probabilities cannot literally be relative frequencies since, among other things, the latter have the wrong formal properties—e.g., they are not countably additive). Facts about relative frequencies of outcomes instead have the status of evidence for claims about the probabilities of those outcomes and it is such facts about probabilities of outcomes (e.g., the probability that a particle with a certain kinetic energy will penetrate a potential barrier of a certain kind) that are explained by statistical theories like QM. The evidential connection between information about relative frequencies and claims about probabilities is established in QM, as elsewhere, by the use of standard statistical tests such as tests of significance—tests whose role would be quite opaque if we did not make something like the above distinction. On this conception one thinks of the explanations provided by theories like QM as having something like an ordinary DN structure or as what Hempel calls deductive-statistical (DS) explanations. What is explained by such theories is (just) what can be deduced from them—claims about probabilities of individual outcomes. So construed, such theories do explain facts about (do apply to) individual outcomes and not just large collections—although what is explained is not why such particular outcomes occur, but rather why they occur with a certain probability.

I suspect that one reason Salmon does not take this possibility as seriously as he might is that the paradigms of statistical theorizing with which he works involve very low-level phenomenological generalizations about the behavior of particular kinds of systems, e.g., generalizations about the half-life of particular radioactive isotopes. If one thinks that generalizations of this sort are typical of the generalizations which figure in the explanans of microphysical explanations—that they (perhaps when supplemented in the appropriate way with information about continuous causal processes) are what do the explaining—it will be hard to avoid the conclusion that if anything is explained by microphysical explanations, it is the occurrence of particular outcomes. If the information that a carbon-14 atom has a probability of 1/2 of emitting an electron within a period of 5730 years and that such atom is a continuous causal process explains anything at all (as Salmon supposes (46–47, 202–4)), what could it explain but why such an emission will occur on a particular occasion?

However, although paradigms of this sort have dominated philosophical dis-

cussions of statistical explanation, it seems to me doubtful that they are good examples of the sorts of explanations provided by a statistical theory like QM. It is much more natural and in accord with scientific practice to think that explanation of the behavior of some system in QM characteristically involves the solution of the Schrödinger equation for that system given facts about the Hamiltonian governing the system and other facts about initial and boundary conditions. On this sort of conception, the sorts of low-level probabilistic generalizations about the behavior of particular kinds of systems described above are among the explananda and not part of the explanans in typical quantum-mechanical explanations, for it is such generalizations that one derives in solving the Schrödinger equation for a particular kind of system. Explanation in QM does not come in when, e.g., one subsumes some particular episode of barrier penetration under a generalization that tells us that whenever a particle of such and such kinetic energy encounters a potential barrier of such and such a shape, it has such and such a determinate probability of tunneling through. (Nor does it come from seeing such an episode as the result of a continuous causal process.) Rather, one possesses an explanation when facts like those expressed in the above generalization are derived in an appropriate way from a generalization like Schrödinger's equation of much wider scope and when one comes to see how this derivation is an instance of a much wider pattern of derivation, in which Schrödinger's equation is solved with respect to a variety of different microphysical systems. On this alternative conception QM would not explain anything—it would not be a serious candidate for a physical theory at all—if it just consisted of a vast collection of generalizations about such matters as the half lives of various kinds of atoms, the behavior of electrons in particular kinds of potential wells, and so forth. Understanding these particular kinds of probabilistic behavior as part of a much more general pattern is essential to the sort of understanding QM provides.¹⁷

Clearly, drawing attention to this alternative picture of what a statistical theory like QM explains does not by itself show that Salmon's own views are misguided. However, I think it does suggest that it is unlikely that Salmon's conception can be supported just via appeal to what philosophers find it reasonable or intuitive to say about specific examples of the use of statistical theories (after all, these "intuitions" have been formed largely by reading the philosophical literature on the subject). Such examples—including those Salmon appeals to in *SE*—can always be reconstrued as ones in which what is explained is a probability or else denied to be cases of statistical explanation at all.¹⁸ Instead, it seems to me that the best way to approach the question of what statistical theories explain is to ask whether there is any real work that is done by Salmon's model which could not be done by the alternative conception elaborated above.

Consider, for example, the problem of assessing the explanatory power of rival quantum mechanical theories—say Bohr's early theory of 1913 vs. elementary nonrelativistic QM in its modern formulation, or the latter theory vs. quantum

electrodynamics. Are there any considerations that are relevant to assessing the explanatory power of these theories which would be left out if one were to reject Salmon's conception of statistical explanation in favor of the alternative DS conception considered above? I think that much of the appeal of the DS conception derives from the suspicion that the answer to this question is "no" and that the sorts of considerations which are relevant to the assessment of the explanatory power of statistical theories are just the sorts of considerations that are relevant to assessing the explanatory power of deterministic theories – familiar considerations having to do with what can be derived from the two theories, with how unified, non-ad hoc each theory is, and so forth. Thus, for example, one might say, within the framework of a purely DS account of statistical explanation, that Bohr's quantum theory provided a better explanation of spectral emissions of hydrogen (where it at least made accurate predictions) than it did of nonclassical barrier penetration (which it did not predict at all). Again, one might say that Bohr's theory does not explain very well, if at all, why an electron in a potential well will occupy only discrete energy levels (Bohr's "quantum conditions" are imposed ad hoc, without any real justification besides the fact that they yield experimentally correct results – another nice example, incidentally, of arbitrariness in explanation), while modern quantum mechanics provides a much better explanation of this phenomenon (the quantization of allowable energy levels arises in a natural way – as the solution to an eigenvalue problem – out of the imposition of certain boundary conditions on Schrödinger's equation).

In making comparisons of this sort, we do not seem to need to appeal to the notion of statistical explanation of individual outcomes. Adherents of the DS conception will think that this is true generally and that there is nothing about our practices of methodological assessment of how well real-life statistical theories explain which requires the introduction of such a notion. If this is so, and if, as Salmon's discussion leads one to suppose, any plausible conception of statistical explanation of particular outcomes will have counterintuitive features (e.g., that E_1 will explain both E_2 and not E_2), this seems to me to be a good reason to try to get along without such a conception. Conversely, to defend his claim that statistical theories like QM explain individual outcomes, it seems to me that Salmon must show that we need his model for purposes of methodological assessment in connection with serious, realistic examples – that, say, we need to appeal to a model of the statistical explanation of particular outcomes if we are to make plausible comparisons of the explanatory power of the different quantum-mechanical theories described above.

IV

Finally, I want to conclude by commenting on Salmon's remarks on logic-oriented or epistemic conceptions of explanation. While much of what Salmon

has to say on these topics is interesting and convincing, I thought that his discussion suffered from (a) a tendency to run together a number of distinct questions and (b) especially in connection with his discussion of the SR basis, a failure to fully come to terms with the apparent implications of these remarks for his own model of explanation. With regard to (a), one might well want to distinguish the following questions: (i) Can one determine on a priori grounds, and quite independently of contingent facts about the world (from logic alone), what sort of criteria a good scientific explanation must satisfy? (ii) Even if one cannot do this, can one discover, as a result of an a posteriori investigation of features possessed by paradigms of good explanation in our world, necessary and/or sufficient conditions statable in purely syntactic or semantic terms for any purported explanation (regardless of subject matter) to be acceptable? (iii) Even if the answers to (i) and (ii) are negative, do good explanations often possess a fair amount of explicit deductive structure and should one think of them as explaining in part in virtue of providing such deductive structure? Does whether or not we have an explanation of some phenomenon (e.g., the behavior of some gas or of barrier penetration) have something to do with whether one knows how to write down and solve an equation associated with that phenomenon rather than just providing looser specification of an underlying mechanism without such an explicit derivation? (iv) Should one in trying to characterize the notion of explanation assign a central role to such characteristically "epistemic" or "inferential" notions as diminishing arbitrariness or achieving unification?

It seems to me that Salmon is surely right in answering (i) and (ii) in the negative, and that it is an important insight on his part that many previous accounts of explanation are defective because they imply an affirmative answer to (i) and (ii). However, it does not follow from the rejection of (i) and (ii), (and is arguably false that) that the correct answer to (iii) and (iv) is "no." One can share Salmon's skepticism regarding content-free, subject matter-neutral, purely "logical" characterizations of explanation and agree with his suggestion that whether a certain explanatory strategy or criterion for explanatory goodness is likely to be reasonable or fruitful is heavily dependent on various empirical facts about the domain to which it is to be applied and yet continue to believe that explanation is in important respects an epistemic notion and that in many but not all areas of investigation (physics, but not history) whether one has a good explanation has a lot to do with whether (and what sort of) derivation one has.¹⁹

With regard to (b), Salmon's general claims about the contingent character of explanation raise an obvious question concerning the range of applicability of the causal/mechanical model. Although Salmon's discussion is nondogmatic in tone and is tempered by various qualifications, it also exhibits a clear tendency to suppose that the causal/mechanical model will apply to most domains of scientific investigation, at least outside of quantum mechanics.²⁰ But if one gives up the expectation that a model of explanation should be universal in the sense that it must

apply to all logically possible worlds, why should one continue to expect that such a model must be universal in the sense that it applies to all or even most domains of investigation in our world? Given the general point that which explanatory principles we will find “useful and appealing” will depend on general facts about the causal structure of our world (240), why not take the further step of admitting that different explanatory principles may be useful and appealing in, say, physics, psychology, and sociology, depending on general, contingent facts about the characteristic subject matters of these disciplines? Taking this step would allow one to recognize that the causal/mechanical model captures important features of explanation in classical physics and yet to deny that the model is likely to be illuminating in connection with, say, information-processing explanations in psychology. It would also allow one to resist the temptation to say that physical theories like elementary quantum mechanics are unexplanatory to the extent that they fail to specify continuous causal processes.

Salmon’s discussion of the SR basis seems to me to represent a striking illustration of this general point. Although, as we have noted, the SR basis is plainly intended to be applicable to examples in QM such as radioactive decay, it is also extensively illustrated by reference to examples from other sciences such as sociology and epidemiology. Indeed Salmon’s first and only fully worked out illustration of the SR basis involves a sociological explanation of juvenile delinquency in terms of such factors as class and religious background and parents’ marital status (37–43). Here too, it seems to me that—quite apart from my critical remarks about the appropriateness of thinking of quantum mechanics as providing explanations of individual outcomes in section III above—Salmon is much too willing to assume that there is some single, unitary notion of statistical explanation or some single way in which statistical information is relevant as a first stage in the construction of explanations (as in “the” SR basis), which all of his examples illustrate.

In sociological cases (and often in epidemiological cases as well) researchers typically employ one or another so-called causal modeling technique—e.g., regression analysis or path analysis. Here it seems to me one really is interested in something resembling (what Salmon calls) “statistical relevance relations,” although in scientific practice such relations are expressed in the form of data regarding variances and covariances and the correlation and regression coefficients calculable from such information and not in the form prescribed in Salmon’s SR basis.²¹ When one uses such techniques one is, in effect, making inferences about causal connections on the basis of information about statistical relevance relations in conjunction with certain other nonstatistical information just as Salmon’s discussion of the SR basis suggests.²² Like the SR stage itself, causal modeling techniques are, in important respects, data-driven or inductivist procedures for theory construction. They are used when one has a great deal of statistical information about the incidence of various factors of interest in specific popu-

lations, but lacks a precise, predictively powerful theory of the sort found in many areas of physics and chemistry. From statistical information, including information about statistical relevance relations as expressed in claims about covariances and various other assumptions, one infers, say, values for the coefficients in a regression equation, which are taken to reflect facts about structural causal connections.

Furthermore, causal modeling techniques possess a number of other features at least roughly resembling those Salmon assigns to the SR basis. For example, as the SR basis requires, it is natural to think, in connection with such techniques, that one begins with a reference class which is specified by the population from which one is sampling (American teen-agers). One will accordingly have statistical data specifying a determinate "prior probability" (or at least prior frequency) of occurrence of the explanandum variable (e.g., juvenile delinquency) in that class. It is then natural to think, as Salmon's discussion suggests, of explanation in terms of the introduction of further variables (class, religious background, and so forth) which "partition" this original reference class and which are relevant to the likelihood of incidence of juvenile delinquency. The conclusion one reaches, just as in the case of the SR basis, will be relative to the specific population with which one begins – it will be, at best, a conclusion about the causes of delinquency among contemporary American, but not contemporary Chinese, teen-agers.

Now contrast the above features with the sorts of explanations provided by a statistical theory like QM. The statistical information which is used to test QM is not what one would naturally think of as information about statistical relevance relations between the dependent and independent variables in the theory (that is, information about covariances or correlations) at all, but is rather in effect information about the frequency distribution of measured variables (position, momentum, etc.) which can be used to test the probabilistic predictions of the theory. QM is a precise, integrated, predictively powerful theory of a kind one typically does not have available in the sort of context in which one uses regression analysis. One would not expect (or need) to infer the values of key parameters in such a theory, reflecting claims about causal connections, from the data as in the case of regression analysis; rather the theory itself prescribes values for such parameters. It would be mad to try to construct, arrive at, or confirm QM by running regression analyses on, say, statistical information about the incidence of radioactive decay in various populations of atoms (or on spectroscopic data or on any other body of information which constitutes the evidential basis for QM). Moreover, the claims of QM are not specific to particular populations of microphysical systems, as is characteristic of the SR basis, but are intended to apply universally to all systems (anywhere in the universe) possessing certain very general features. It is *not* natural to think of explanation in QM in terms of beginning with information about the incidence of an explanandum variable in

some specific reference class and then partitioning this reference class into subclasses on the basis of information about statistical relevance relations.²³

If this were the end of the matter, it would be reasonable to conclude that the SR basis captures important aspects of the use of causal modeling techniques, but not (even as a “first stage”) of explanation in quantum mechanics. But other features that Salmon assigns to the SR basis seem not to fit at all with the use of causal modeling techniques, although they do in some cases correspond to features of explanation in quantum mechanics. Consider, for example, the requirement that the explanans variables in the SR basis effect an objectively homogeneous partition—roughly that no further statistically relevant subdivision (meeting certain other natural conditions) of the partition established by the explanans variables be possible. The various no-hidden variable results perhaps ensure that (something like) this requirement is satisfied in quantum mechanics. On the other hand, it seems to be generally accepted that such a requirement will virtually never be satisfied in the sorts of contexts in which causal modeling techniques are typically used. Given a regression equation containing certain variables, it will always be possible to find other variables the inclusion of which in the model will change previous claims about statistical relevance relations (in this case, the values of the original coefficients in the regression equation).²⁴

One can see this clearly in the case of Salmon’s own example concerning juvenile delinquency. Given a partition created by explanans variables having to do with class, religious background, and so forth, it looks as though there will be an indefinite number of further variables (finer gradations in personal income, child-rearing practices, geographical location, schooling, characteristics of the criminal justice system, and so forth) which will be statistically relevant. This fact certainly raises interesting methodological problems regarding the use of causal modeling techniques and points to the need for restrictions on the class of candidates for explanans variables when such techniques are employed. But to impose the requirement that any explanation of a phenomenon like juvenile delinquency must employ a partition such that no further statistically relevant subdivisions are possible is in all likelihood to ensure that no explanation of this phenomenon will ever be forthcoming. Instead, what one wants is an account of explanation which makes it clear how the citing of general explanans variables like class can be explanatory, even though further relevant partition is plainly possible.

Finally, we should note another important disanalogy between Salmon’s SR basis and techniques like regression analysis. As noted above, Salmon thinks that statistical information, whether in quantum mechanics or sociology, is used to explain individual outcomes. Thus, in connection with his juvenile delinquency example, he suggests that statistical information about the incidence of juvenile delinquency can be used to explain why some particular boy, Albert, commits a delinquent act (37). But even if one wants to think of a statistical theory like QM as explaining individual outcomes, it seems doubtful that this is the correct way

to think about the juvenile delinquency example. When one uses a technique like regression analysis (or more generally assembles information about statistical relations for the purposes of constructing an explanation), it seems more plausible to suppose that one is trying to explain facts about population level parameters—facts about changes in the mean incidence of juvenile delinquency in the American population or perhaps facts about the variance of this variable. One is interested in trying to explain such facts as the great increase in juvenile crime in the 1960s and 1970s, or why, on the average, youths from urban areas commit more crimes than youths from rural areas, and not in explaining why any particular boy became a juvenile delinquent.

I thus suggest that the kind of explanations provided by QM and by the use of causal modeling techniques (which represent the way in which information about statistical relevance relations is used to make causal inferences in actual scientific practice) differ in a number of quite fundamental respects, which in turn reflect differences in subject matter, in available information, and so forth. In my view, Salmon's SR basis represents an attempt to join together in one unified account conditions on explanation drawn from several quite different ways in which statistical consideration can figure in the construction of explanations. Many of these conditions correspond to quite real features of either explanations in QM or of explanations based on statistical relevance relations, but there is no single, unitary notion of statistical explanation (or of the evidential basis for such explanations) which combines all of these disparate conditions. Once one abandons the demand for a single, universal logic of explanation it seems to me to be equally natural to abandon the idea that there must be some single model of statistical explanation possessing all of the features Salmon assigns to the SR basis.

Although I have been critical of a number of Salmon's claims, I hope it is clear from my discussion that *SE* is a valuable and provocative book. It provides a clear and perhaps definitive statement of a distinctive conception of scientific explanation, a conception that unquestionably captures central aspects of causal explanation in many physical contexts, even if it is less generally applicable than Salmon supposes. And it develops a number of general themes—the role of individual causal processes, the limitations of purely formal models, the importance of contingent facts about how nature actually works in the characterization of explanation—that one suspects will be at the center of philosophical discussion for some time to come.

Notes

1. Salmon apparently regards these various features of causal processes as co-extensive but this is not obviously correct. An electron, for example, when represented quantum mechanically as a wave packet can be marked and transmits a determinate structure and a determinate probability of being found in various states on measurement, but there are certainly limits on the extent to which it can be appropriately regarded as a spatio-temporally continuous process, with a definite location. Similarly consider a macroscopic sample of gas. This can transmit a mark and, when in a certain state

of temperature and volume, transmits a determinate probability of exerting a certain pressure. If our standards are not too fine-grained, it can even be regarded as a spatio-temporally continuous process. Yet, as I note immediately below, Salmon denies that an explanation of the pressure of the gas in terms of the ideal gas law (or, one presumes, an explanation which appeals to the fact that the macroscopic sample is a causal process with a determinate probability of exerting a certain pressure) satisfies the constraints of the causal/mechanical model. Here the idea seems to be that a satisfactory explanation must cite more fundamental microphysical mechanisms underlying the behavior of the gas, or perhaps explicitly cite processes which involve the local transfer of energy and momentum, as a statistical mechanical analysis would do. The various criteria on which Salmon seems to rely in these cases for determining whether an explanation appropriately cites causal processes do not necessarily coincide. An explanation may cite processes which transmit structure (which are not pseudo-processes) without detailing fundamental mechanisms or explicitly detailing mechanisms which are mechanical in the sense that they involve the transfer of energy and momentum. Indeed, on pp. 202–3 of *SE* several examples (including one in which a subject in a psychological experiment is said to have a determinate probability to respond in a certain way) are given of processes which transmit structure but which do not cite, at least in any detail, any fundamental mechanisms and these are apparently regarded by Salmon as acceptable explanations (or at least as examples of causal processes or the transmission of causal influence).

I think that more needs to be said by Salmon about the interrelations among these various features of causal processes and about which features should be taken as criterial. Why, for example, do the ideal gas laws fail to cite causal processes, while the above description of the psychological experiment apparently does? Why not take the ability to transmit a mark as the fundamental feature of a causal process (whether or not such transmission satisfies a spatio-temporal continuity constraint) and thus take the evolution of the wave function in QM to represent a legitimate causal process?

2. A factor C is statistically relevant to the occurrence of B in reference class A if and only if $P(B/A \cdot C) \neq P(B/A)$. I should add that this is just the bare bones of the SR basis; for the full characterization, which is much more complex, see *SE*, pp. 36–37.

3. It is a further question, however, whether it is appropriate to think of the notion of a spatio-temporally continuous causal process as a primitive notion for purposes of explanation, even in contexts in which the notion applies. Salmon seems to hold, or at least to suggest, especially in his discussion of explanation in quantum mechanics (see below) that there is just something intrinsically more intelligible about a process that involves local rather than remote conservation or action at a distance. Processes of the former sort are, as it were, natural stopping places in explanation. My own quite different view (and also, I think, the view of many physicists) is that the existence of spatio-temporally continuous causal processes is itself something that stands in need of explanation—one can perfectly intelligibly ask *why* the world should have this feature.

One natural suggestion is that the explanation for this fact is, as already indicated, bound up with Special Relativity and with the fact that the structure of spacetime is locally Minkowskian. This alternative conception fits naturally with the idea, urged briefly below, that a more fundamental element in our conception of a good physical explanation is that such explanations should satisfy certain symmetry and invariance (e.g., a requirement of Lorentz invariance). It is because we think that theories satisfying such requirements provide especially good explanations that we are led, derivatively, to value explanations containing continuous causal chains and local transfers of energy and momentum. A somewhat similar view is defended by Brian Skyrms in his *Causal Necessity*, (1980) pp. 110–27, and by Hans Ohanian in his *Gravitation and Spacetime* (1976). Ohanian writes, a passage quoted by Skyrms, that

Nowadays, all the fundamental interactions are regarded as due to local fields. . . . Why do we prefer fields to action-at-a-distance? The answer is simple: we need fields in order to uphold the law of conservation of energy and momentum [given the requirement of Lorentz invariance]. (36).

4. For a defense of the idea that the ascription of geometrical structure (particularly affine and metrical structure) to spacetime is explanatory, see Graham Nerlich's *The Shape of Space* (1976) and Michael Friedman's *Foundations of Space-Time Theories* (1983), and for an explicit defense of the idea that such explanations are noncausal, see Nerlich. That explanation in General Relativity does not seem to fit the causal/mechanical model very well is also noted by Clark Glymour in his "Causal Inference and Causal Explanation" (1982).

I might also note that the question of whether the ascription of structure to spacetime is explanatory is presumably closely bound up with the adequacy of various "causal" theories of the structure of space time and with whether various "conventionalist" or anti-realist theses regarding spacetime are correct. Someone who thinks that claims about spacetime structure, if taken literally or realistically, must be reducible to claims about actual and perhaps possible causal processes and interactions and who thinks that the ascription of space-time structure which is not so reducible is a matter of convention, will presumably deny that such structure can figure in explanations. If the structure in question is conventional, it can hardly figure in explanations, and if it is not conventional, and is reducible to talk about causal processes, then on a view like Salmon's, one should presumably just appeal to facts about such processes in constructing explanations. The idea that the ascription of spacetime structure can be explanatory in its own right seems to require realism about such structure and the rejection of the claim that such structure is reducible to causal relations.

Since Salmon has defended conventionalist views elsewhere and has explicitly claimed that the attribution of affine and metric structure in General Relativity is a matter of convention (1977), I strongly suspect that that he would just deny that General Relativity explains, via the postulation of geometrical structure, in the sense I have claimed. My own view is that in view of the widespread rejection of conventionalism and of purely causal theories of spacetime (at least in conjunction with General Relativity) in recent philosophical work, the apparent philosophical connection between the causal/mechanical conception of explanation and causal theories of space-time structure represents a liability of the former conception. At any event, the connection deserves further examination. Both are expressions of the idea that the fundamental furniture of the universe is (just) causal processes and their interactions. I suspect that both stand or fall together.

5. At least I take this to be the tendency of Salmon's discussion, which is tentative in character and perhaps does not reach an unequivocal conclusion. (Salmon tells us, on the final page of *SE*, that he has "not offered any account of quantum mechanical explanation and I do not believe that anybody else has done so either." (279)). I should also acknowledge that while Salmon seems to regard quantum mechanics, at least on its presently available interpretations, as not fully satisfactory from the point of view of furnishing explanation or understanding, he certainly recognizes the possibility adumbrated below: that instead of revising microphysics to bring it more into accord with the requirements of the causal/mechanical model, perhaps we ought to revise the requirements on explanation embodied in the causal/mechanical model and change our views about the fundamental mechanisms in the world on the basis of the apparent microphysical facts. Thus, for example, he writes at one point:

Does this kind of conservation [that is the kind of "remote" conservation apparently involved in the EPR correlations] require explanation by means of some special mechanism? Or is this one of the fundamental mechanisms by which nature operates? These questions strike me as profound and I make no pretense of having an answer to them. (258)

6. I have no general account of when it is reasonable to conclude that the demand for explanation of some general kind is misguided or inappropriate. This will depend upon a variety of complex empirical and conceptual considerations and must be determined on a case-by-case basis. But plausible illustrations are not hard to find: for example, demands for explanations which appeal to essences or essential properties or to final causes are certainly inappropriate in many, if not all, domains. It is not a limitation on the explanatory power of current physical or chemical theories that they do not provide answers to all the kinds of explanatory questions Aristotle would have regarded as appropriate. More controversially, it is arguably no limitation on the explanatory power of theories in neurobiology or psychology that they provide no explanation of what it is like to be a bat, in the sense intended by Thomas Nagel in his well-known paper (1974) of that title.

7. Salmon rather closely follows Bernard d'Espagnat's recent discussion (1979) which seems to adopt this assumption. Thus Salmon says that the measurement of a component of spin of one particle in the EPR experiment "alters" (251) the spin component of the correlated particle and appears to endorse d'Espagnat's suggestion that "nature seems to exhibit action at a distance in the quantum domain" (250). But the standard view, in any event, is that such talk of "altering" or of action at a distance in the sense of the transmission of causal influence or information makes sense only if the correlated particle has a definite spin component prior to measurement. It is precisely by rejecting this assumption that the standard view attempts to avoid interpreting quantum phenomena as involving action at

a distance or violations of "local" causality. Moreover, it is well known that if one adopts the contrary, realistic assumption, one is led to the conclusion that various quantum phenomena like barrier penetration actually violate conservation laws, a conclusion Salmon would presumably wish to avoid.

8. See, for example, Bell (1964), Kochen and Specker (1967), Gleason (1957), Aspect and Roger (1982). The connection between those results and Salmon's discussion is perhaps not as clear as one would like—in part because there is disagreement about the interpretation of the results themselves, and about just what they rule out, but also because Salmon himself is not very explicit about what he thinks a common-cause explanation of the EPR correlations would involve. For example, there are a number of results like Bell's theorem which rely on various versions of a locality condition to derive inequalities in contradiction with the statistics predicted by quantum mechanics and actually obtained in experiment. In versions of Bell's theorem which apply to deterministic local hidden variable theories, the locality condition is often interpreted as (or motivated by reference to) the special relativistic requirement that faster than light signals be impossible (but see Earman [1986] for the contrary view that the locality condition required for Bell's theorem has nothing to do with Einstein locality and instead represents semantic locality). Moreover, if the correlations in EPR type experiments are regarded as obtaining strictly and universally, one can show that the local hidden variable theory producing this result must be deterministic. On the other hand, if one is willing to regard the strict obtaining of the correlations as an idealization, which holds only approximately in the real world, then one can consider stochastic hidden variable theories. To derive a Bell-like inequality for such theories, one must assume an (apparently) stronger locality condition, the motivation and physical significance of which is a source of some controversy (see, e.g., Hellman [1982], Jarrett [1984]). As Jarrett shows, this condition has a natural decomposition into relativistic locality and an additional condition that he calls completeness and that is also sometimes called factorizability of conditional stochastic independence—roughly that measurement results on the two correlated particles be independent, conditional on the values of the hidden variables.

Now Salmon gives no indication that he supposes that conservation laws hold only approximately in EPR type correlations and seems to accept that such correlations do not involve superluminal action (*SE*, 250). More generally, his conception of a common cause explanation seems to be such that it satisfies some fairly strong version of a locality condition: the idea is that a common cause operates via a local interaction and the subsequent propagation of the results of the interaction via distinct continuous causal processes to the correlated measurement results, each process containing the "instructions" relevant to the result to which it is connected. One would think that this sort of conception rules out, say, a purported explanation of the outcome of performing a measurement on particle L in which the characterization of the common cause requires reference to the measurement outcome obtained for the correlated particle R or reference to the experimental arrangement employed at R. (Compare van Fraassen (1982) for arguments showing how one can derive a Bell type inequality given reasonable restrictions of this kind on what a common-cause explanation for the EPR correlations would have to look like.) Given all of this, it is very hard to see how there would be any sort of common-cause explanation of the EPR correlations meeting other conditions Salmon would accept.

On the other hand, if sufficiently weak restrictions are placed on what may count as a common cause—if a stochastic, contextual hidden variable theory which doesn't satisfy a strong locality constraint will qualify, then of course it will be possible to reproduce the observed correlations. It would have been very useful to have had a more explicit discussion of all this in *SE*. Just what is the relationship between the demand for a common-cause mechanism and results like Bell's? In which, if any, of the several different senses of locality must a common-cause explanation satisfy a locality requirement? Which theories in the space of possible hidden variable theories (deterministic vs. stochastic, local vs. nonlocal, contextual vs. noncontextual) would, if empirically adequate, count for Salmon as providing common-cause explanations?

9. For relevant see Elliot Sober's "Equilibrium Explanation" (1983).

10. For example, one prominent high-energy physicist comments that it is a "strength" of gauge theories such as a GCD in comparison with earlier theories that "they require comparatively few free . . . parameters." Nonetheless

even if the free parameters have been reduced to a manageable number, they remain an essential part of the theory. No explanation can be offered of why they assume the values they do. (t'Hooft 1980, 136)

By contrast an attraction of recent superstring theories is that they lack such arbitrariness—such theories apparently can only be consistently formulated in 10 dimensions and with one of two gauge symmetry groups. They thus yield the hope of explaining why spacetime has the number of dimensions it does and why nature prefers the symmetry groups it does. As a recent discussion in *Physics Today* puts it:

Unlike the (nongravitational) grand-unified point-field theories (GUTS) of recent years, where the gauge symmetry groups and coupling parameters could be chosen with considerable freedom and fit the data, the superstring theories offer almost no free choices. (Schwarzschild 1985)

The correspondent in *Science* expresses a similar idea

[in superstring theory, the] symmetry groups were defined by the underlying model, instead of being adjusted by hand to fit the data. For the first time there seemed to be a mechanism [although of course not a mechanism in Salmon's sense—JW] for nature to choose her symmetry group. (10)

11. The idea that theoretical unification plays a central role in scientific explanation is defended by, for example, Michael Friedman, (1974), Philip Kitcher (1981), and James Woodward (1979).

12. See, for example, Stephen Weinberg "The Search for Unity: Notes Toward a History of Quantum Field Theory" (1977).

13. This general point is made very clearly and elegantly by Philip Kitcher in the context of a discussion of the relation between Mendelian genetics and molecular biology in his "1953 and All That: A Tale of Two Sciences" (1984). One of Kitcher's illustrations is particularly apt: Mendelian genetics accounts for various facts about gene transmission in part in terms of meiosis and the independent assortment of genes on nonhomologous chromosomes. However, the molecular processes which underlie meiosis in different organisms are quite heterogeneous. An account which just traced such detail in the case of a particular species would lose the more general pattern embodied in Mendel's laws.

14. The claim that complex systems exhibit levels of organization and that explanation often proceeds by abstracting from certain kinds of lower-level causal detail and finding general patterns does not imply that one can simply ignore constraints owing to lower-level causal facts in attempting to discover such general patterns. How much two systems can offer in causally relevant detail at some lower level of analysis and still fruitfully be regarded as relevantly similar at a higher level of analysis and how much lower-level features constrain possible upper-level patterns is not something that can be stipulated a priori but requires detailed empirical inquiry on a case-by-case basis. Philosophical defenses of functionalist doctrines in psychology notoriously ignore this point.

15. Of course it is perfectly true that a particular gene, organism, or psychological subject is a continuous causal process, but presumably it does not follow *just* from this observation that the above explanations successfully explain by citing (or just by citing) continuous causal processes and interactions. Or, if this conclusion does follow, it is not clear what sorts of explanations the causal/mechanical model rules out.

16. Another reason (regarding which I comment below) has to do with Salmon's failure to distinguish between statistical theories like QM and the use of information about statistical relevance relations to make causal inferences, and his tendency to think that techniques of the latter sort are used to establish claims about individual causal connections.

17. Salmon's intuition here is, as I understand it, a very different one: it is that (if I may put it this way) all that there fundamentally is in the world is particular atoms with particular determinate propensities to decay, particular electrons, particular electromagnetic fields, and so forth. On an "ontic" conception of explanation it must be such facts about particular causal processes and interactions which constitute the raw material out of which our explanations are constructed—it is these that are really most central or significant for purposes of explanation. The significance of a high-level generalization like Schrödinger's equation is, if I have understood Salmon's view correctly, derivative from or parasitic on these facts about particular causal processes—the equation is, at best, an abstract description of what lots of individual processes have in common.

18. Salmon does (117–18) describe a number of concrete scientific examples (including the Davisson-Germer electron diffraction experiment, Rutherford's scattering experiments, Compton scattering, and an example from genetics) which he claims must be understood as cases in which a statistical theory is used to explain an individual outcome. While I lack the space to discuss these in detail, I found this claim unpersuasive, both for the general reasons given in the main body of this paper and because all the examples are quite underdescribed and, in a number of cases, do not clearly

involve subsumption under a statistical generalization at all. For example, Salmon does not spell out what he has in mind when he speaks of providing a theoretical explanation of the Rutherford scattering or Davisson-Germer diffraction results, but the usual elementary textbook treatments (and the treatments originally given by Rutherford and Davisson-Germer themselves) are essentially classical and appeal to fundamental laws which are deterministic rather than statistical. In the Davisson-Germer experiment the analysis makes use of the relation $\lambda = h/p$ but otherwise proceeds along the same lines as ordinary optical diffraction and results in the derivation of an expression (the Bragg relation: $n\lambda = 2d \cos \theta$) which gives the intensity maxima in the reflected beam. It is the existence of these maxima which seem to be explained via the derivation, and the derivation seems to have a straightforward DN structure. In the case of Rutherford scattering, in the usual treatment one assumes that the scattering is due to a repulsive coulomb force and derives, from classical considerations, the equation of motion for an incident particle and from this an expression for the differential scattering cross-section (or for the fraction of particles scattered at a given angle). Here again the derivation looks like a straightforward DN derivation which involves no subsumption under irreducibly statistical laws. A similar point holds for Compton scattering: here one uses the conservation of energy and momentum and the relation $p = h/\lambda$ to derive an expression ($\lambda' - \lambda = \frac{h}{m_0c} (1 - \cos \theta)$) for the shift in wave-

length of the reflected photon as a function of the scattering angle θ . (See, e.g., Anderson [1971] for relevant discussion.) It is perfectly true that in a full quantum mechanical treatment of scattering that one represents both the incident and reflected flux of particles as wave functions and assumes that they must obey the Schrödinger equation, so that the treatment is, in this sense, irreducibly statistical. But here again what seems to be explained is just what can be derived—e.g., facts about the scattering cross-section or about the probability of finding a particle (or the particle flux) at a given scattering angle. At the very least, I think that Salmon needs to spell out his examples in much more detail and make it clear just exactly what in his view is being explained, what is doing the explaining, what sort of theoretical treatment is envisioned, and so forth. Just because the processes involved in scattering experiments are “irreducibly statistical” it does not follow that the explanations we give for features of those processes must be statistical explanations, in Salmon’s sense.

19. While it is plausible that there are no necessary and perfectly general or sufficient conditions statable in purely syntactic or semantic terms for explanatory goodness, it is quite compatible with this that various features of good explanations in specific domains can be given interesting formal (but domain-specific) characterizations which fall short of specifying necessary and sufficient conditions. For example, we have already noted that theoretical unification represents an important goal of physical explanation. While there is arguably no perfectly general formal characterization of explanatory unification, a natural way of partially characterizing such unification in the context of fundamental physical theories is via the imposition of symmetry and invariance requirements. I am inclined to think that there may be other domain-specific formal criteria for explanatory goodness in other areas of investigation—see, for example, Clark Glymour’s “Causal Inference and Causal Explanation” (1982) and his “Explanation and Realism,” (1984) for some interesting suggestions along this line appropriate to causal modeling techniques.

20. At several points in his discussion, Salmon acknowledges the possibility that the causal/mechanical model may not apply everywhere in our world. Thus he writes: “I make no claim for universal applicability of my characterization of scientific explanation in all domains of our world, let alone for universality across all possible worlds” (240). At a number of other points, however, he reveals expectations that are considerably more universalist regarding explanation in our world: “I hope that the causal theory of scientific explanation outlined above in this book is reasonably adequate for the characterization of explanation in most scientific contexts—in the physical, biological and social sciences—as long as we do not become involved in quantum mechanics.” (278)

21. For helpful discussion of connections between Salmon’s notions of statistical relevance and screening off and various aspects of the use of causal modeling techniques see Suppes (1970) and especially Irzik (1986). The analogy seems closest in connection with causal models with dichotomous (two-valued) variables and in connection with procedures like those developed by Herbert Simon and Herbert Blalock for choosing among causal models on the basis of information about correlation coefficients. As one moves to more general linear models with continuous variables, the analogy becomes much more attenuated if not positively misleading. In these models the relevant notion of cau-

sation seems to be the notion of counterfactual supporting functional dependence, and to have little to do with positive statistical relevance or the tracing of causal processes in Salmon's sense. For additional discussion, see Woodward (1988).

22. There is, however, an important disanalogy between Salmon's treatment and causal modeling techniques at this point, which deserves a more detailed exploration than I can give it here. Both Salmon's treatment and causal modeling techniques reject the idea that one can infer claims about causal connections just from facts about patterns of statistical association. Both agree that additional information is required to support such an inference. However, they disagree in an important way about the character of the additional information which is required. For Salmon, the additional information has to do centrally with facts about individual causal processes and their interactions and thus with facts about temporal order and spatio-temporal continuity. By contrast, while such considerations certainly sometimes play a role in the sorts of contexts in which causal modeling techniques are used, they are often not of central or decisive importance. Thus, although one's conclusions about whether two variables of interest are causally related can be greatly affected by which additional variables one includes or excludes from a regression equation (see note 24), considerations of spatio-temporal continuity are often not very useful in considering which variables to exclude and often are not appealed to by users of regression equations. For example, whether or not a variable measuring the execution rate seems to causally affect a variable measuring the murder rate may depend on whether one includes variables reflecting the employment rate and poverty level in the regression equation. But considerations regarding spatio-temporal continuity or demands for the tracing of continuous causal processes are often simply not helpful in deciding whether it is reasonable to include these variables—instead users of causal modeling techniques typically appeal to general theoretical considerations bearing on whether the poverty level is a "possible cause" of changes in the murder rate.

Similarly, given the sort of data one often has available when one uses causal modeling techniques (e.g., cross-sectional data at a relatively high level of aggregation) and the nature of the models themselves which are not dynamical in character or precise about temporal relations between variables, it is often impossible to use information about temporal order to disambiguate the direction of causality. Given two correlated variables X and Y, arguments about which is the cause and which is the effect instead will turn on general theoretical or common-sensical claims about whether X is the sort of thing which could cause Y or vice versa. Thus in studies of voting behavior it is commonly (and I think falsely) assumed that a voter's judgement that a political candidate has views similar to his own can cause a decision to vote for that candidate but not vice versa.

Extra-statistical assumptions of these sorts about possible causes or about causal order are plainly already causal in character, and so causal modeling techniques are emphatically not techniques for deriving causal claims from purely noncausal premises; they are rather techniques which show us how we can test causal claims on the basis of statistical information if we are willing to make other causal assumptions. To the extent that Salmon's treatment involves the idea that purely statistical information (the SR basis) when supplemented just with considerations of spatio-temporal order and continuity will allow us to sort out genuine causal connections from noncausal sorts of association, it is more "Humean" and reductionist in spirit than techniques like regression analysis. It is also at odds with the actual practice of those who use such techniques.

23. Consider, for instance, Salmon's example of explaining the decay of a carbon-14 atom. What should we understand the reference class A to be here? If we make what might seem to be the natural choice of taking A to be the class of all carbon-14 atoms, then it appears that no further statistically relevant partition of this reference class is possible—there are no further factors that are statistically relevant to the decay and so the basic requirement of constructing the SR basis by introducing a further partition is not met. To even get started constructing the SR basis, we need to construe the demand for explanation as relative to some reference class we can partition—say, the class of all carbon atoms in a certain sample. If we have information about the prior frequency of decay in this reference class, we can then partition the reference class into, e.g., those atoms that are radioactive carbon atoms and those that are stable isotopes—this will be a further statically relevant factor and will thus form part of the SR basis and part of the basis for any explanation we construct.

But even as a "first-stage" account of explanation in QM (or of the role of statistical evidence in constructing such explanations) this seems misdirected. The prior probability of decay in the above

reference class will depend upon quite idiosyncratic and contingent facts about the frequency with which carbon-14 and stable carbon atoms occur in the sample, and as we choose different reference classes or different samples, this probability will vary greatly. (I take it that something like this point underlies Nancy Cartwright's claim that the S-R model involves an undesirable kind of relativization to a reference class in her [1979]). The sorts of explanations of decay and related phenomena found in quantum mechanics textbooks do not seem to invoke the above information about prior probabilities (indeed in many cases this probability will be entirely unknown) and give no indication of involving this sort of relativization to a reference class. If our aim is the characterization of how explanation works in QM, it is not plausible that if someone asks why this particular carbon-14 atom decayed, we can give no explanation (or assemble no statistical information bearing on the construction of an explanation), but if someone asks why this carbon atom from a mixed sample decayed (and we happen to know the prior probability of decay and this happens to be different from the posteriori probability), we do have the basis for constructing an explanation and the explanation we give is that it is a carbon-14 atom with such and such a probability of decay. The explanation we give for the behavior of the atom in question should be the familiar quantum mechanical one involving the penetration of a charged particle by a potential barrier and it should be the same in both cases. While some causal explanations genuinely do involve relativization to a reference class and proceed in effect by partitioning it—this is characteristic of many explanations produced by causal modeling techniques—this sort of conception just does not seem to fit QM.

24. Suppose we are trying to estimate the coefficients in the general linear regression equation

$$Y = B_1X_1 + B_2X_2 + \dots + B_nX_n.$$

The basic point is that the partial regression coefficient B_i , which (supposedly) tells us about the causal or functional connection between the dependent variable Y and the independent variable X_i , depends not just on the covariance between X_i and Y but also on the covariance between X_i and the other independent variables in the equation. If, say, we originally do the regression just using variable X_1 and then add X_2 to the regression equation, the coefficient for X_1 will change as long as X_1 and X_2 exhibit a non-zero correlation in the data. Given any finite body of data—and it is worth emphasizing that of course this is the only kind of data we ever actually have access to—it is very likely that we are always going to be able to find such additional, correlated variables. Regression analysis thus always requires additional extra statistical assumptions about which variables to include in the regression equation—assumptions which will draw on prior theoretical ideas about causal and explanatory relevance. The idea, perhaps suggested by Salmon's treatment of the role of the SR basis, that one should proceed by first assembling "all" information about statistical relevance relations, without causal or theoretical presuppositions, and only then trying to construct a causal explanation on the basis of this information, is not something it would be sensible (or even possible) to carry out in practice.

References

- Anderson, Elmer. 1971. *Modern Physics and Quantum Mechanics*. Philadelphia: W. B. Saunders.
- Aspect, A., Dalibard, J., and Roger, D. 1982. Experimental Test of Bell's Inequalities Using Variable Analyzers. *Physical Review Letters* 49:1804–7.
- Bell, J. S. 1964. On the Einstein-Podolsky-Rosen Paradox. *Physics I*: 195–200.
- . 1966. On the Problem of Hidden Variables in Quantum Mechanics. *Reviews of Modern Physics* 38:447–75.
- Cartwright, Nancy. 1979. Causal Laws and Effective Strategies. *Noûs* 13:419–37.
- D'Espagnat, Bernard. 1979. The Quantum Theory and Reality. *Scientific American* 241, no. 5 (November):158–81.
- Earman, John. 1986. *A Primer on Determinism*. Dordrecht, Holland: D. Reidel.
- Friedman, Michael. 1974. Explanation and Scientific Understanding. *Journal of Philosophy* 71:5–19.
- . 1983. *Foundations of Space-Time Theories*. Princeton, NJ: Princeton University Press.
- Gleason, A. M. 1957. Measures on the Closed Subspaces of a Hilbert Space. *Journal of Mathematics and Mechanics* 6:885–93.

- Glymour, Clark. 1982. Causal Inference and Causal Explanation. In *What? Where? When? Why?* ed. Robert McLaughlin. Dordrecht, Holland: D. Reidel.
- . 1984. Explanation and Realism. In *Scientific Realism*, ed. Jarret Leplin. Berkeley: University of California Press.
- Hellman, Geoffrey. 1982. Stochastic Einstein-Locality and the Bell Theorems. *Synthese* 53:461–504.
- Hempel, Carl G. 1962. Deductive-Nonlogical vs Statistical Explanation. In *Minnesota Studies in the Philosophy of Science*, Vol. 3, eds. Herbert Feigl and Grover Maxwell. Minneapolis: University of Minnesota Press.
- . 1968. Maximal Specificity and Law Likeness in Probabilistic Explanation. *Philosophy of Science* 35:116–33.
- Irzik, Guro. 1986. Causal Modeling and Statistical Analysis of Causation. *PSA*, Vol. I. East Lansing, MI: Philosophy of Science Association.
- Jarrett, Jon. 1984. On the Physical Significance of the Locality Conditions in the Bell Arguments. *Nous* 18:569–89.
- Kitcher, Philip. 1981. Explanatory Unification. *Philosophy of Science* 48:507–31.
- . 1984. 1953 and All That: A Tale of Two Sciences. *Philosophical Review* 93, no. 3, 335–73.
- Kochen, S., and Specker E. P. 1967. The Problem of Hidden Variables in Quantum Mechanics. *Journal of Mathematics and Mechanics* 17:59–87.
- Nagel, Thomas. 1974. What Is It Like to be a Bat? *Philosophical Review* 83, No. 4, 435–50.
- Nerlich, Graham. 1976. *The Shape of Space*. Cambridge: the University Press.
- Ohanian, Hans. 1976. *Gravitation and Space Time*. New York: W. W. Norton.
- Salmon, Wesley. 1971. Statistical Explanation and Statistical Relevance. In *Statistical Explanation and Statistical Relevance*, ed. Wesley Salmon. Pittsburgh: University of Pittsburgh Press.
- . 1977. The Curvature of Physical Space. In *Minnesota Studies in the Philosophy of Science*, Vol. VIII, *Foundation of Space-Time Theories*, eds. John Earman, Clark Glymour, and John Stachel. Minnesota: University of Minnesota Press.
- Schwarzschild, Bertram. 1985. Anomaly Cancellation Launches Superstring Bandwagon. *Physics Today* 38, No. 7, 17–20.
- Skyrms, Brian. 1980. *Causal Necessity*. New Haven, Ct: Yale University Press.
- Sober, Elliot. 1983. Equilibrium Explanation. *Philosophical Studies* 43:201–10.
- Suppes, Patrick. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.
- t'Hooft, Gerard. 1980. Gauge Theories of the Forces between Elementary Particles. *Scientific American* 242, no. 6 (June):104–38.
- van Fraassen, Bas. 1982. The Charybdis of Realism: Epistemological Implications of Bell's Inequality. *Synthese* 52:25–38.
- Weinberg, Stephen. 1977. The Search for Unity: Notes Toward a History of Quantum Field Theory. *Daedalus* 106, no. 4, 17–35.
- Woodward, James. 1979. Scientific Explanation. *The British Journal for the Philosophy of Science* 30:41–67.
- . 1988. Understanding Regression in *PSA 1988*. East Lansing, MI: Philosophy of Science Association.