Theories of Causation and Explanation: Necessarily True or Domain-Specific?*

Paul Humphreys
512 Cabell Hall
P.O. Box 400780
University of Virginia
Charlottesville, VA 22904
pwh2a@virginia.edu

with Comments by†

Julian Reiss
Centre for Philosophy of Natural and Social Science
London School of Economics
Houghton St
London WC2A 2AE
j.reiss@lse.ac.uk

July 2004

 I should like to thank Julian Reiss for his comments on this paper, and various members of the audience at the "Causality and Explanation: Homage to W. Salmon" conference, Barcelona, August 4-5, 2003, for their reactions to the original, rather different, version.
† I would like to thank the Arts and Humanities Research Board for supporting this research.

1 Introduction

In the course of his journey from the statistical relevance model of explanation through the causal account to the conserved quantities theory, Wes Salmon was responsible for a number of significant changes in our way of thinking about both explanation and causation. Perhaps the most profound contribution that he made to the literature on these subjects was his insistence that ontic accounts of both causation and explanation can be made plausible. This kind of scientific realism, Salmon insisted, had to be compatible with a recognizable form of empiricism. And in so doing, he seemed to make the task of constructing and evaluating a satisfactory theory of explanation much more difficult than it had hitherto been. For unlike most previous accounts of explanation and causation, Salmon's ontic accounts seem to be based on empirical facts rather than on conceptual analysis and they are, if true, contingently true. One of Salmon's great insights was to see that we must understand the world on its own terms rather than to insist that a single *a priori* explanatory framework is adequate. This is exactly what a scientific realist should require. If our world had been different and devoid of causal processes then we should have to explain things differently. Salmon's realism is, nevertheless, slanted in a particular direction. I argued in Humphreys [2000] that because they are based on conserved quantities, Salmon's final theories of causation and explanation require a commitment to a quite radical form of physicalism, making the theory difficult to apply to social, economic, and psychological phenomena This commitment would seem to narrow the scope of the theories considerably unless a comprehensive reduction of the social and psychological sciences to physics can be achieved.

I now think that the empirical content inherent in Salmon's theories, although still present, is less problematical than I once thought and this makes the philosophical problems involved in evaluating them less severe than they might appear. Furthermore, the degree of commitment to physicalism that is required in Salmon's most mature theories is quite minimal. Taken together, these suggest that the combination of realism and empiricism in Salmon's theories requires less of a change in philosophical methods than might be expected.

2 The Modal Status of Theories of Causation and Explanation.

There are five kinds of features that theories of causation and, separately, theories of scientific explanation, may possess. The most familiar is that they are necessarily true[1], where the modality is metaphysical necessity, *i.e*. what the theory says about causes and explanations is true in all possible worlds. This is the approach standardly taken by those wishing to define causation, as for example, David Lewis did.[2] So, if you subscribe to a counterfactual theory of causation, it is not an accident that causes happen in our world to be such that if the cause event had not occurred, then the effect event would not have occurred either. It is in the very nature of a causal relation that this must be true. Whatever the world might be like, wherever there is a causal relation, it must have the properties stipulated in the definition. This is a familiar property of any account arrived at by conceptual analysis. As a second, albeit less interesting, possibility, a

[1] I shall assume that all theories I discuss are true in order to avoid the continual use of 'if true, then necessarily true' and related hypotheticals.

[2] Lewis [1973], [1986]. Because the definitive version of Lewis's later paper 'Causation as Influence' has not been published at the time of writing, I have not classified that theory.

theory of causation or explanation might involve conceptual necessity, *i.e*. it is impossible for cognitive agents similar to ourselves to consistently conceive of a counter-example to the theory. What is the strength of the impossibility in this claim? That depends upon your theory of mind. My own view, because I hold that psychological properties emerge from biochemical and physical properties, is that the appropriate impossibility for this position is a nomological impossibility, but others, especially those who adhere to a particular kind of supervenience position, may be able to make a case that it involves a metaphysical impossibility. Conceptual necessity is weaker than metaphysical necessity, for it allows that there might be worlds, the causal structure of which is conceptually unintelligible to us, within which causes operate but in ways we can never grasp.

As a third possibility, the theories of causation and explanation might appeal to nomological necessity, *i.e*. the theory of causation or explanation is true in all worlds having the same scientific laws as ours but there might be other kinds of causation in worlds with different laws. The fact that nomological possibilities can be involved in both the second and the third kinds of possibility does not make the second kind a special case of the third. To hold that it is nomologically possible or impossible for a cognitive agent to consistently conceive of a counter-example to a theory is quite different from asserting that such a nomological possibility does or does not exist. It is a striking feature of minds that they are capable, despite being physically implemented, of transcending, within the conceptual realm, what is nomologically possible.

These three kinds of necessity can be contrasted with a fourth feature that might be possessed by a theory of causation or explanation, that of being universally true, independently of subject matter. Although there need be no modal content to this kind of theory, it still makes a striking claim – that there is a uniform set of features possessed by all causes. It asserts, amongst other things, that there is no kind of causation special to physics or special to sociology.

The last kind of causal or explanatory theory is one that is domain specific or subject matter specific, one that is true for some subject matters but inapplicable to others. For a theory to be subject matter specific is for its conditions of adequacy either to make essential reference to some specific subject matter or subject matters or to preclude the application of those conditions to certain kinds of subject matter.

These divisions are important because they affect how we assess the truth of these theories. Traditional *a priori* analyses can effectively address the truth status of metaphysically or conceptually necessary theories of causation and explanation, but each of the last three kinds of theory requires additional, empirical, knowledge. The more domain specific the theory, the more scientific and the less philosophical the theory appears to be. This is not an insurmountable objection to domain specific theories, but it does require us to say something explicit about how they are to be judged.

3 The Status of Salmon's Theory

The two principal divisions used in the taxonomy above, the more traditional one of necessary versus contingent theories, and the perhaps less familiar one of universal versus domain specific theories, are such that all four combinations are logically possible. There can be necessarily true domain specific theories and contingently true universal theories, as well as contingently true domain specific theories and necessarily true universal theories. Most traditional analyses of causation are intended to be both universal and necessarily true. In contrast, if the structure of

our world is such that an ontological reduction of the subject matter of all other sciences to that of physics is possible, but this reduction is a contingent fact about our world, then a theory of causation that grounds causation in physical processes will give a contingently true but universal theory. As a third possibility, because David Fair's energy transmission theory is nomologically necessary and, if the kind of reduction just discussed is not possible because the laws of certain non-physical sciences are sui generis, then Fair's will be a nomologically necessary but domain specific theory.[3]

My main focus in this paper is the status of Salmon's theories and in order to assess his final accounts of causation and explanation, we need to trace the intellectual history that led to their development. Salmon's early statistical-relevance approach to explanation (Salmon [1970]) undercut the nomic expectability basis of Hempel's deductive-nomological and inductive statistical models of explanation by showing that events with low probabilities could be explained. This was done by citing factors that assigned events to objectively homogenous reference classes. To explain an event was to assign it to the broadest such homogenous reference class, that class being arrived at by successive specifications of statistically relevant factors. Because such a theory needs to distinguish between statistical associations that are the result of direct causal connections and those that are not, Salmon began to construct a theory of probabilistic causation. Yet the task of completing Hans Reichenbach's purely probabilistic theory of causation proved elusive for Salmon. Faced with the problem of distinguishing between mere correlations and genuine causal relations, Salmon decided to view statistical relevance relations as simply the evidential starting point of explanations, a basis that itself had to be explained. This was done in terms of an ontology of spatially continuous processes and of interactions between those processes. The criterion of a genuine causal process, as opposed to a pseudo-process, was its ability to have its structure altered after an interaction. This criterion seemed to have a counterfactual claim embedded in it – what would happen if a mark were to be introduced – and conscientious empiricists such as Salmon wanted no part of irreducible counterfactual claims. I should add that it was not just empiricist qualms that led Salmon to reject counterfactuals but a dissatisfaction with their context-dependence: 'A major part of the motivation for [the change from a marked processes view to a conserved quantities view] was an aversion to counterfactuals. I was seeking completely objective causal concepts; counterfactuals are notoriously context dependent.' (Salmon [1997], p. 470). So following the lead of Phil Dowe (Dowe [1992a]), causation was instead taken to consist in the transfer of conserved quantities: '...causal processes transmit conserved quantities; and by virtue of this fact, they are causal...' (Salmon [1994], p. 303).

Salmon's account of explanation based on the conserved quantity theory of causation is a straightforward example of a domain specific contingent theory. Salmon was quite explicit about this when he wrote: 'What constitutes an adequate explanation depends crucially, I think, on the kind of world in which we live; moreover what constitutes an adequate explanation may differ from one domain to another in the actual world. ... The ontic conception mandates attention to the mechanisms that actually operate in the domain in which explanation is sought' (Salmon [1985] p. 299). Salmon also maintained that ontic explanations are usually grounded in causes: 'According to the ontic conception – as I see it at least – an explanation of an event involves

[3] I am not aware of any metaphysically necessary domain specific theories of causation or explanation but such things are surely possible. In mathematics, for example, arithmetic and geometry are necessarily true but domain specific.

exhibiting that event as it is embedded in its causal network and/or displaying its internal causal structure.' (Salmon [1985], p. 298), but he also allows that causal explanations are a subclass of the more general class of ontic explanations: '...causal explanations of the sort just discussed are adequate and appropriate in many domains of science but that other mechanisms – possibly of a radically non-causal sort – operate in the quantum domain'[4] (Salmon [1985], p. 298).

This lack of universality is in stark contrast to the situation with Hempel's deductive-nomological and inductive-statistical models of explanation because the deductive-nomological and inductive-statistical models of scientific explanation are both metaphysically necessary. (The logical empiricists, of course, would have been horrified by that terminology.) They are necessary and also universal because their central concept of nomic expectability based upon logical inference is applicable independently of subject matter. For those subjects or worlds lacking laws – a famous and controversial candidate was history – there simply are no Hempelian explanations: 'The decisive requirement for every sound explanation remains that it subsume the explanandum under general laws.'[5] But what of other worlds in which there exists subject matter not of our world – suppose, to take an elementary case, that a genetically engineered species had been invented that fitted into none of the biological categories existing in our world? Although Hempel did not as far as I know address this kind of issue, it is reasonable to infer from the way his theories were presented that nomic expectability was essential to any adequate scientific explanation and that such other-worldly cases must conform to his criteria of adequacy for explanations.[6]

Of the other rivals to Salmon's theory in the area of explanation, the pragmatic account's context-dependency is not inherently subject matter dependent. The relevance relation changes from context to context because of a questioner's interests but while these interests may differ with areas of investigation, they can equally well be determined by factors that are independent of the subject matter. As for the unification approaches, it has often been pointed out that whether or not they are universally true in our world is a contingent matter, but that objection underestimates the degree to which Kantian motivations are driving the theory. Indeed, the whole thrust of the unification approach is to erase the boundaries between areas that are currently considered to be separate.

In contrast to his later conserved quantity account, Salmon's original statistical relevance account was universal because the statistical relevance relations upon which the theory is based are subject matter independent. Whether a given factor A is statistically relevant to the frequency of another factor B within reference class R is, of course, a contingent matter, discoverable only by examining sequences of empirically generated data, but the statistical relevance relations themselves, at least on the relative frequency approach, rest on arithmetical relations and hence can be applied whatever the origin of the data.

I mentioned earlier that Salmon's shift to a conserved quantity account seems to commit him to physicalism in a fairly dramatic way. There are few, if any, conservation laws in the social sciences and the emphasis on conserved quantities thus seems to either severely limit the scope of the theory or to commit him to a reductionist or eliminativist program of a quite extreme kind. It seems to entail that all anthropological and sociological causation, for example, must be

[4] Salmon goes on to say in the next paragraph that there could be non-causal ontic explanations and that there may be quantum mechanical explanations for which continuous causal processes are not involved.
[5] The quotation is from p. 258 of the slightly revised 1965 reprint of Hempel [1948].
[6] This is clear in the case of D-N explanations from Hempel's four conditions of adequacy for that model (Hempel [1948], pp. 137-138).

accounted for in terms of the transfer of mass-energy, linear and angular momentum, and other conserved quantities. And so all explanations in those sciences must ultimately be given in terms of physical causation. Perhaps this is what physicalists believe, and it presents in a stark form the kind of reductionist ontology that produces the various problems of mental causation. There is, in consequence, a kind of ineliminable subject-dependence in the conserved quantity account of causation, a dependence upon the availability of a certain kind of physics underlying all phenomena we consider to be causal, and there is no reason to think that this dependence must hold across all subject matters in all possible worlds, whether those worlds are nomologically or metaphysically possible. In contrast, analyses in terms of sufficient or necessary conditions, or those in terms of statistical relevance relations, have no difficulty in dealing with social or economic causation because they are universal accounts and in consequence can remain neutral on the reductionism issue. Even Salmon's middle period theory of markable processes can, in principle, be applied to any material subject matter. Marked bills, distorted states of consciousness, changes in social structure, are all plausible examples within, respectively, economics, psychology, and sociology.

4 Hitchcock's Criticisms and the Hybrid Theory

So, Salmon's theories of causation and explanation are neither necessarily true nor universal. Despite this, I think that Salmon's theory is much more general than it might appear from the features discussed in the last section. A considerable amount of emphasis is placed by Salmon on what he calls causal processes and causal interactions. Yet there is in his use of the term 'causal' the possibility for a serious confusion, a possibility that is compounded by not infrequent references to the transmission of causal influences. We can see that the role played by the term 'causal' in these references to 'causal processes' is minimal by recalling that more than one conserved quantity can be transmitted by a causal process and that more than one conserved quantity can be exchanged during a causal interaction. If it were the transmission of 'causal influence' that was crucial for causation, we would have the analog of a causal overdetermination problem for the conserved quantity theory. As it is, either conserved quantity will do to make the processes and interactions causal and it is irrelevant which quantity is conserved or exchanged as long as at least one is present. And indeed, the explanatory factors in Salmon's final theory are not ordinarily those that make the interactions and processes causal. Under the pressure of criticisms due to Chris Hitchcock[7], Salmon modified his theory in the mid-1990s so that it became a hybrid. The appeal to conserved quantities distinguishes between causal and pseudo-processes, but the explanatory factors in any given case must include those factors that are statistically relevant. As Salmon put it: 'In [*Scientific Explanation and the Causal Structure of the World*] I characterized scientific explanation as a two tiered structure, consisting of statistical relevance relations on the one hand and causal processes and interactions on the other. As a result of Hitchcock's analysis, I would now say (1) that statistical relevance relations, in the absence of connecting causal processes, lack explanatory import and (2) that connecting causal processes, in the absence of statistical relevance relations also lack explanatory import. In various discussions I have focused on (1) to the virtual neglect of (2)...this was a mistake. Both are indispensable' (Salmon [1997], p. 476). The re-emphasis on the statistically relevant factors in the conserved quantity approach thus restores many of the examples that made the statistical-

[7] See Hitchcock [1995].

relevance model initially plausible. We explain delinquency in terms of broken homes, unemployed fathers, depressed economies, and low levels of education, not in terms of conserved quantities. Of course, mass-energy conservation holds in the social realm as well, but it is not the relevant variable.

This means that despite the frequent use of examples involving colliding billiard balls, baseballs breaking windows, spots of light moving along walls, and so forth, the role played by the so-called causal processes in explanations is minimal. All that is required is a commitment to a widely held form of physicalism, the view that every non-physical property is carried by one of these causal processes, or, as I prefer to call them, 'carrier processes'. Thus, every economic mechanism, every sociological variable, and every psychological property must be conveyed by a physical carrier process. And whether you are a reductionist, a supervenience advocate, a property dualist, or almost anything other than a substance dualist or an idealist, this is a small commitment indeed. It does not get us any degree of necessity for the processes and interactions part of Salmon's theory, although I believe that one could argue that the nomological necessity of the ontic view ought to follow from this position if conservation laws were insisted upon rather than mere regularities. In addition, the fact that the principal explanatory role is played by statistically relevant factors means that a considerable part of, although certainly not all, the explanatory theory is indeed subject-matter independent.

5 A Residual Tension

Yet there remains something unsatisfactory about this hybrid account. There had long been a tension in Salmon's theory between, on the one hand, the kind of example typified by the hexed salt example, where what was important was the relevancy or irrelevancy of cited characteristics and, on the other hand, the kind of considerations which were introduced by examples such as the transmission of heritable characteristics and the decay of radioactive atoms, for which it was not the increase or decrease of the probability value that was important but the transmission of a probability value or distribution. The rejection of the increase in probability approach to explanation was, it seems safe to say, motivated for Salmon by examples such as those involving the transmission of genetic characteristics that he frequently cited in support of the transmission of probabilities view. And in fact such appeals to statistical mechanisms underlay the entire ontic account of causal processes: 'Scientific understanding according to [the ontic] conception involves laying bare the mechanisms – etiological or constitutive, causal or non-causal – that bring about the fact to be explained'(Salmon [1985], p. 301). On the ontic view, explanation must promote understanding of how the world works and that understanding is mechanical: In a previously unpublished article that first appeared in his *Causation and Explanation* collection (Salmon [1998]), Salmon wrote: 'I shall examine two general forms of scientific understanding...The second involves understanding the basic mechanisms that operate in our world, that is, knowing how things work. This kind of understanding is mechanical.' (Salmon [1998a], p. 81) Salmon goes on to say a few pages later 'It is the kind of understanding we achieve when we take apart an old-fashioned watch, with springs and cogged wheels, and successfully put it together again, seeing how each part functions in relation to all the others' (Salmon [1998a], p. 87).

In order to lessen the tension between the statistical-relevance view and the transmission of probabilities view, I want to suggest an approach that lies within the broad outlines of the

empirical realist framework Salmon has given us. It is, I am sure, one with which he would have disagreed. So I shall present the arguments for the alternative, but leave the conclusion in the form of a choice that must be made between two quite different ways of conceiving of causation and explanation. Each has its merits and although I have a preference for one of them, I can quite well see the appeal of the other.

6 Conserved Quantities Are Not Sufficient For Causation

We saw above that in his 1997 article (Salmon [1997]), Salmon had suggested that a greater degree of balance needed to be achieved between the emphasis on causal processes and mechanisms and the emphasis on statistically relevant factors. To provide an ontic explanation is to specify the mechanisms and patterns that were involved in the production of the explanandum, and these mechanisms are to be analyzed in terms of the process account. There are two aspects to this process approach adopted by Salmon. One is the 'at-at' formulation of processes, which conforms to the idea that there are no mysterious non-humean connections between events within single processes. The other is his rejection of regularity accounts of causation. In a 1985 article he wrote: 'It may be possible – thought I seriously doubt it – to construct a regularity analysis of causality that would be adequate within the context of Laplacean determinism... Although I do not have any knock-down argument to support my contention, my sense of the objections to [theories of probabilistic causality based on statistical regularities] convinces me (at least tentatively) that no such regularity analysis of probabilistic causality will be adequate. We must instead look to mechanisms' (Salmon [1985], pp.296-297) Let me begin with the at-at theory as it is employed in the conserved quantities account of causation.

Salmon's formulation of the conserved quantities theory is encapsulated in the following three propositions:

1. A causal interaction is an intersection of worldlines which involves exchange of a conserved quantity. (Salmon [1994], p. 303)

2. A causal process is a worldline of an object that transmits a nonzero amount of an invariant quantity at each moment of its history (each space-time point of its trajectory). (Salmon [1994], p. 308)

3. A process transmits a conserved quantity between A and B ($A \neq B$) if and only if it possesses [a fixed amount of] this quantity at A and at B and at every stage of the process between A and B without any interactions in the open interval (A, B) that involve an exchange of that particular conserved quantity. (Salmon [1997], p. 462)

And he says of this definition that '...it yields a criterion that is impeccably empirical, and thus it provides an acceptable answer to the fundamental problem Hume raised about causality' (Salmon [1997], p. 469).

This definition is perhaps *too* humean, for proposition 1 allows as causal interactions things that are clearly neither causal nor interactive. Consider two marching bands wearing identical uniforms. They are skilled at performing the kinds of intricate pass-through maneuvers characteristic of such bands but they lack a certain sort of discipline. Whenever the bands intersect, the individual members spontaneously decide whether to reel off left or right according

to their own whim. They are sufficiently skilled that nobody ever collides with another bandsman, and the emerging bands, which will be composed of mixtures of members of the original bands and will ordinarily be of different sizes, always coalesce into two ordered bands ready for the next intersection. Now take L = total number of bandsmen in the left hand band and R = total number of bandsmen in the right hand band. L + R is a conserved quantity--the total number of bandsmen emerging from a band intersection is always the same as the number going in. This renders the above definition of a causal interaction applicable and the two marching bands constitute two causal processes involved in a causal interaction simply because of the conservation of number principle This should strike you as odd.

In response to this objection, Salmon responded in conversation that number is not an conserved quantity – for example, placing two rabbits of the opposite sex in a hutch will quickly result in a serious violation of rabbit number conservation. Point taken, but there are three counters to this reply. First, if this kind of response were legitimate it would show that the laws of arithmetic were not necessarily true, and there are well-known strategies one can adopt in order to show that such examples are misplaced. A less trivial counter is to note that Salmon allows, in response to a point made by Phil Dowe (Dowe [1992b]), that he, Salmon, wants the theory of causal processes and interactions to be considered at the theoretical level, where a number of idealizations and abstractions from actual processes is permissible. (Salmon [1997], p. 464). This is a perfectly reasonable move to make and we can preserve the conservation of number principle by imposing closure conditions on rabbit systems. These will be straightforward constraints on rabbit (and bandsmen) interactions – no opposite sex interactions, no applicability to pregnant rabbits, confinement to a closed region, and application to living rabbits only. Indeed, with these restrictions on the scope of the generalization, conservation of rabbit number is as good a candidate for a true conserved quantity principle as are the genuine scientific laws of conservation of baryon number and conservation of lepton number. To remind you what these are: baryons are particles which exert strong nuclear forces and have fractional spins. Examples are neutrons and protons. In all known interactions, the number of baryons entering into an interaction is equal to the number of baryons exiting the interaction. A similar law holds for leptons, which are particles exerting weak interactive forces and having spin 1/2, such as electrons, positrons, muons, and neutrinos. Finally, it might be said that the conservation of number principle is not a scientific law. If that were so, then it is sufficient to point out that Salmon does not require generalizations about conserved quantities to be law like; he merely requires that they be true ([1994], p. 310). He does this to prevent re-entering the modal circle via law likeness, a re-entry that he wants to avoid, given that the motivation for shifting from the mark transmission criterion for causal processes to the transmission of invariant quantities was to avoid recourse to counterfactual criteria for mark transmission. But if all that is required is for the invariance of number principle to be true, rather than to be a law, then with the idealizations mentioned earlier, that condition is satisfied.

So, we have an example that seems to show Salmon's theory is too broad. I mentioned above that it is odd to call something an interaction where what we have is, in the case of the marching bands, merely a spatial coincidence and concomitant change. But its oddness is merely a reflection of what is involved in these humean accounts of causal interactions. In the marching bands case there is no interaction of any recognizable kind--the bandsmen reach their appointed spot on the field, then scatter in one of two directions by a purely chance-like process. In fact the scattering could be arranged by spontaneous fusion in a single band after a set period of time (thus creating a ‘y fork’) and the interactive content of this is even less obvious.

This detour through the nature of causal interactions has brought us back to the basic issue of the modal status of Salmon's account. For one thing that these considerations suggest is that some conservation laws play the role of what Michael Friedman calls 'constitutive *a priori* principles'.[8] Roughly speaking, constitutive *a priori* principles are principles that are not metaphysically necessarily but must be adopted in order for a specific scientific theory to be applied. Well-known examples of constitutive *a priori* principles are the choice of geometry in classical gravitational theory and general relativity, and the adoption of Newton's three laws in classical mechanics. Certain conservation principles such as the conservation of energy are good candidates for the role of constitutive *a priori* principles. If they are, then because such principles are *a priori* in nature – they are accepted or rejected on grounds other than coherence with or conflict with empirical data – that aspect of the ontic account has something like a philosophical rather than a scientific aspect.

That said, I shall finish this section by noting a residual problem. Unless we can exclude the conservation of number principle somehow – and note that it is has empirical content in virtue of the subject-specific idealizations that are required to make it true for a given domain – the conserved quantities approach does not even mandate the kind of minimal commitment to physicalism that I described earlier. For even granting Quine's concerns about the difficulties of counting certain non-physical entities such as possibilia and beliefs, there is no difficulty at all in counting objects that are specifically social or economic in form. There are currently thirty three fraternities at the University of Virginia. A fraternity is not a physical entity – it is partly a social unit, partly a cultural unit, partly a legal entity, and it exists quasi-independently of its current human members and physical assets. It thus seems possible to apply certain conservation principles to social entities without any commitment in those cases to physicalism. Deciding what to do with such examples leads us into interesting but complex metaphysical territory and so I shall not pursue the issue further here.

7 Conclusion

What does all this say about Salmon's causal theory? I mentioned that Salmon's most recent account of causation and explanation was a hybrid theory. It contains references to processes and interactions on the one hand and to statistically relevant factors on the other. How best to weight these components is, I think, a key question that Salmon has bequeathed to us. My own taste, because I hold that causation is primarily a relation between properties (see Humphreys [2004], section 2.9), leans towards emphasizing the relevance relations and a more radical version of the anti-regularity approach to causation than I think Salmon would have been comfortable with. In particular, singular causal relations are basic for me. Yet there is also an undeniable appeal to the kind of neo-humean position that Salmon developed in his at-at account of processes and interactions. It requires us to bring into close contact answers to Why-questions and answers to How-questions and in doing so to rest content with a scientific description of the world. But even if one goes in that direction, I hope to have convinced you that this does not require a wholesale abandonment, but only a modification, of traditional philosophical methods for deciding which approach is the better of the two.

[8] See, for example, Friedman [2001]. The idea, happily, goes back to Reichenbach, who called them 'coordinative definitions'. For a different view of the role played by these principles, see the Appendix to Chapter One of Ryckman [forthcoming].

Comments by Julian Reiss

1 Introduction

In his paper "Theories of Causation and Explanation: Necessarily True or Domain-Specific?", among other things, Paul Humphreys investigates the modal status of Wesley Salmon's theories of causation and explanation and attempts to show that although his later theories are, if true, merely contingently true, a stronger interpretation is possible and (at least, so he implies) desirable.

In my comments I want to do four things. I first inquire into plausible motivations for searching for a stronger than contingent theory of causation and explanation. Second, I show that well-known counterexamples demonstrate that Salmon's theory is not universally true. Third, reflections on philosophical method show that although it is coherent to hold a domain-specific necessary theory, arguing for such a theory would violate Salmon's overall empiricist point of view. Fourth, I very briefly sketch an alternative interpretation of Salmon's theory which does not violate empiricist sentiments.

2 Motivation for Humphreys' Project

In a way, it is a strange question to ask whether a theory of causation and explanation should be contingently or necessarily true. Given how hard it is to find any theory that is even remotely acceptable, not to mention true of a significant number of cases or universally true, why ask further whether it is true only of our world, of any world with the same laws or of all possible worlds?

As far as I can see, there are two plausible candidates for a motivation, one having to do with a philosopher's more general temperament, the other, with philosophical method. Turning to the first candidate, we may exploit an analogy with first-order theories of causation and explanation. One of the dissatisfactions with the covering law model of explanation was that *merely* covering laws in fact appear to explain nothing. In an example due to Nancy Cartwright, the explanation-seeking question is: "Why does the quail in my garden bob its head up and down in that funny way whenever it walks?" The covering law answer, "Because they all do", is unsatisfactory—it is no explanation at all (Cartwright 1983, p. 70).

In a similar way we can say that second-order theories of causation and explanation do not do the explanatory work demanded from them if all they do is to provide a true description of what all cases of causation and explanation have in common. Stating that, say, as a matter of fact all causal interactions *involve* an exchange of a conserved quantity does not tell us anything about what causal relations *are*. To answer the question "Why is this a case of causation?" with "Because they all are like that" fails to give a (second-order) *explanation* at all. Worse, suppose that the alleged theory is not only merely contingently true but also true of only a restricted domain. Now, the answer "Because some are like that" seems to have even less explanatory import.

By contrast, by pointing out that a certain matter of fact instantiates whatever the essence of causation is, we feel we have explained the matter at hand. If, then, a theory is true necessarily and preferably universally, it does provide a satisfactory explanation.

At this point an obvious caution must be inserted. The original reason to develop the

covering law model was the Humean or empiricist anxiety of concepts such as "necessary connection", "hidden powers" and the like. Because concepts or ideas like these have no observable counterparts directly, we had better construct them out of ideas that have observable counterparts. And so we say that what we really mean by necessary connection is regularity plus contiguity plus temporal priority of the cause plus feeling of expectation.

Today we are less worried about providing the sense impression of each and every concept for it to be meaningful. However, empiricists will still demand that we base our theories on the best available evidence. Since evidence can only be gathered in this world, it seems difficult to decide on the basis of evidence, once we have "X causes Y", whether that is true only of this world, of all nomologically identical worlds or of all possible worlds.

The point should be clear now. Especially realistically-inclined philosophers will feel uneasy about "theories" of causation and explanation which are merely contingently true. If, however, the philosopher in search of such a theory accepts certain empiricist strictures, he or she may have no other choice.

The second motivation has to do with philosophical method. The modal status of a theory of causation and explanation will determine what kind of case is admissible as a counterexample to the theory. A metaphysically necessary theory will have to answer any metaphysically possible case, a conceptually necessary theory any logically possible case while a nomologically necessary theory will have to answer only cases that are compatible with our laws of nature and a contingent theory only actual cases. Hence stems a different attitude towards counterexamples. In particular the status of thought experiments is illuminating here. David Lewis, for example, accepted so-called trumping cases as decisive against his own theory of causation although to my knowledge no one has ever provided an actual scientific case of trumping. Trumping is a subclass of a class of cases of what Lewis calls "redundant causation". Cases of redundant causation obtain whenever two or more causes compete to bring about a certain effect. These prove difficult for counterfactual theories of causation because there is always a back-up cause that acts or would act in case the original cause does not or would not, and thus it is not true that the effect would not obtain if the cause were not to. Trumping is a particularly hard nut to crack for counterfactual theories because (a) the two competing causes act to bring about the effect at precisely the same time, (b) they (would) bring about precisely the same effect (at the same time) but (c) only one of the two is eventually responsible for the effect. The original story that introduces trumping involves two wizards (Schaffer 2001, p. 165):

> Imagine that it is a law of magic that the first spell cast on a given day match the enchantment that midnight. Suppose that at noon Merlin casts a spell (the first that day) to turn the prince into a frog, that at 6:00 pm Morgana casts a spell (the only other that day) to turn the prince into a frog, and that at midnight the prince becomes a frog.

This is a problem for counterfactual theories because (*ibid.*):

> Clearly, Merlin's spell (the first that day) is a cause of the prince's becoming a frog and Morgana's is not, because the laws say that the first spells are the consequential ones. Nevertheless, there is no counterfactual dependence… because Morgana's spell is a dependence-breaking backup.

But, one must add, this is a problem only if one accepts uninstantiable thought experiments as genuine counterexamples. Lewis does because his theory is meant to be metaphysically

necessary, and Schaffer's story is surely a metaphysical possibility.[9]

By contrast, Salmon himself (again, to my knowledge) has never used examples that weren't actual or very nearly actual (by that I mean, simplified versions of actual cases). Trumping cases and the like are simply irrelevant for contingent theories of causation and explanation as long as no analogous but actual example can be provided.

3 Humphreys' Interpretation of Salmon's Theory

Humphreys' reading of Salmon's "hybrid" theory of scientific explanation is something like the following (this is my reconstruction of what I take to be Humphreys' interpretation of Salmon's theory):

> *C* explains *E* if and only if
> (a) *C* is statistically relevant to *E*,
> (b) *C* is connected to *E* by a mechanism consisting of causal processes and interactions.

A number of comments are in order at this point (again, in my reading of Humphreys' reading of Salmon). First, (a) and (b) are to be understood in Salmon's reading, that is, (a) *C* raises or changes the probability of *E* in a homogenous class (homogenous with respect to what?) and (b) physical processes and interactions are defined in terms of transmission and exchange of conserved quantities.

Second, according to Humphreys the conserved quantity laws (or some of them) are *a priori* constitutive principles for Salmon's theory in Michael Friedman's sense of "constitutive principle". That is, they are, first, constitutive of other principles in the same theory: without them, the theory's other principles could not be understood, not to mention tested; second, they are *a priori* in the sense that they are not adopted on the basis of generalisation from empirical results. Third, they are only relativised *a priori*, however, in that the constitutive *a priori* principles we do adopt may change and develop with empirical science (Friedman 2001).

The fact that conservation laws are constitutive *a priori* principles for Salmon's theory of explanation in Humphreys' view makes Salmon's theory more than contingently true. Because its basic principles are *a priori*, the truth of the theory of explanation is a philosophical rather than an empirical matter for Humphreys.

Let me mention as an aside that I disagree with this reading of the constitutive *a priori*. "*A priori*" here means "prior to the case at hand", so it is non-empirical for a particular application of the theory but that does not mean that it is not empirical at all. It just means that we need other kinds of evidence to tell for or against it, and maybe that these propositions are more remote than others from data. This, however, does not mean that they cannot be revised in the light of empirical findings. To the contrary, the principles in reflection about which Friedman introduces this terminology were not only revisable but actually revised.

[9] To be fair, the wizard story is not the only case of trumping Schaffer provides. But all other examples I know of (both from Schaffer as well as others) are either pseudo scientific thought experiments (such as Schaffer's case involving fields if different "colours") or everyday cases that are described at such a level of generality that my intuitions run out as to whether they provide genuine cases of trumping. An example of the latter involves two officers of different rank shouting the same command to their troops and, allegedly, the troops obeying only the higher-rank command. I simply don't know whether this is a genuine case of trumping because I have no story as to what's going on in an individual soldier's head when he hears the two commands.

I now want to argue that well-known cases show that neither (a) nor (b) are individually necessary. Against (a) we can cite widely discussed cases where a cause is connected to its effect by well-understood physical mechanisms by it doesn't change it's probability. Examples of such cases include Hesslow's birth control pill's. In the example, a putative cause ("taking birth control") is both negatively and positively relevant to getting thrombosis. By preventing pregnancies, which are themselves a probability-raiser for thrombosis, pills make the incidence of blood clotting less likely. On the other hand, by introducing a certain chemical into the woman's blood stream, they increase the likelihood. Depending on the actual frequencies of the relevant groups in the population, these two effects may just cancel. Hence, probability raising is not necessary for causation. The practical sciences such as medicine and economics are full of examples of this kind because frequently causal relations are exploited in such a way as to insure that effects are not correlated with their causes. A pain killer may contain a substance which causes drowsiness. So the pharmaceutical company that produces it adds a stimulant in order to prevent drowsiness. The pain killer is thus causally related to drowsiness but taking the pain killer may not correlate with getting drowsy.[10] A central bank may aim at stabilising the interest rate by means of its monetary policy tools, say open-market operations. Suppose it achieves its goal. Now the interest rate is constant and thus uncorrelated with any other variable in the system but this is due to the fact that the central bank is able to perfectly causally control it (Hoover 2001, p. 170).

Therefore, statistical relevance may very often be a good indicator of a causal connection but it is not always present and thus cannot be used in a definition of causation.

Social science also provides counterexamples to the necessity of (b). Humphreys argues that "every economic mechanism, every sociological variable, and every psychological process must be possessed by a physical non-pseudo process" in order to make Salmon's theory work. This is what he calls the "minimal commitment" to physicalism. The original point, however, of introducing physical processes was to mark the distinction between correlations that can be accounted for causally and accidental correlations. One class of examples where physical processes might help to get the distinction right involves symptoms of causes that the statistical relevance relation identifies as causes though they are effects. The following example discussed by Salmon in various places is originally due to Ellis Crasnow (Salmon 1997, p. 474, emphasis added):

> A certain businesswoman usually arrives at her office about 9:00 A.M., makes herself a cup of instant coffee, and settles down to read the morning paper before starting her daily work. From time to time, however, she arrives at her office promptly at 8:00 A.M., meets a colleague from another site, and both are served cups of freshly brewed coffee upon arrival. On the mornings when she arrives at 9:00 A.M. she takes the 8:00 bus from home, but when she arrives at 8:00 she takes the 7:00 bus. The taking of the 7:00 bus thus fulfills the statistical conditions (Reichenbach's conjunctive fork) that partly characterize a common cause of the coincidence of the availability of the freshly brewed coffee and the arrival of her colleague. *Catching the earlier bus is not the common cause, however, because appropriate causal connections* [in the process sense] *do not exist*.

Social science cases show that this is exactly what physical processes cannot do for us. Consider the question whether news about firms' profits have an effect on their share prices. How do we test such a hypothesis? In essence, what is done in order to measure the causal effect of news on share prices is to hold fixed all other causes of the share price and take the difference

[10] Nancy Cartwright attributes this example to Lisa Lloyd. See Cartwright 2003, p. 8.

between the average returns on days of profit announcements with "no news" content (*i.e.*, the profit announced is roughly as expected) and the returns on days of profit announcements with news content (*i.e.* the profit announcement either exceeds the prediction ["good news"] or it is lower than the prediction ["bad news"]). If good or bad news, on average, make a difference to the price of the stock, then we regard the hypothesis as confirmed.

The difference between, say, good news and no news events is supposed to be a *causal* difference: in the first case, news has a causal effect on share prices; in the second case news doesn't. But there is no appreciable *physical* difference between the two kinds of events. True, in both cases there are lots of physical processes going on. Consider a case of insider trading. A CFO knows about the (good) news announcement some days in advance. At a romantic dinner, he tells his mistress. She mentions the affair next morning to her husband. He tells his son, and his son gives a hint to his girlfriend. She calls her broker and the broker places the order. All these are, supposedly, physical processes in Salmon's sense. But whether or not the share prices will rise has nothing to do with that. It has to do with whether or not people, on average, interpret the news as good and act accordingly by buying the stock. Nobody will be able to predict by investigating the kind of physical process involved (genuine or pseudo) what will happen to the share price.

To put the matter differently, that physical processes are a necessary ingredient of causal economic relationships in some sense may be true but it is entirely uninteresting. It is really like saying that causality is equal to *X* (say, statistical relevance) plus God exists. Most importantly, the difference between a genuine and a spurious case of causation isn't marked by the fact that we find a connecting physical process in one case, and no such process in the other. All economic relationships (causal as well as spurious) involve myriads of physical processes but these are not (always[11]) what makes them causal or spurious.

Above I pointed out that degree of domain-specificity and modal status are orthogonal properties. We can have highly domain-specific but necessary theories as well as universal theories that are only contingently true. As far as I am aware, transference theories have always been advanced as universal but contingent theories.[12] Now that we have seen that neither probability raising nor physical processes are essential to causation, it seems we need to give up on the aim of providing a universally true theory.

But we still might be able to find domain-specific theories which are more than contingently true. This gets us back to my claim made at the beginning about the significance of philosophical method for Humphreys' project. If one's aim is to show that a theory of *X* is necessarily true, it does not suffice to point out that it agrees with all known cases of *X*. It has to be shown that all those cases agree with the theory as a matter of metaphysical or conceptual or, possibly, nomological fact (this division is of course Humphreys'). It seems to me that this could be done if necessity is understood in a Kripkean sense. Suppose that billiard ball *A* causes ball *B* to move. Suppose further that it is true that in so doing *A* imparted a certain amount of the conserved quantity momentum on *B*. Now it doesn't seem entirely implausible to say that in this case it is essential to *A*'s causing *B* that momentum was transferred. To the contrary, taking momentum away will result in taking the causing away.

This might not always be the case, however. We can also imagine more complex cases of

[11] Of course there may be non-sense correlations where the absence of a physical connection gives us a reason to believe that the correlation is not genuine. Elliott Sober's oft-cited British bread prices and Venetian sea levels may constitute such an example (see Sober 2001).

[12] At any rate, they have been interpreted as such. See for example Tooley 1990, p. 216.

causation where if one route of transmission of causal influence is suppressed, another one is conceivable. But I do not want to take stance in the debate whether if *A* causes *B* in a particular way in a particular case, *A* does so of necessity or it could have achieved it in a different way. The point I do want to make is about philosophical method. The traditional method to evaluate modal claims is philosophical analysis and the preferred tool is the thought experiment. Imagine a causing and suppose away the transfer of conserved quantity (or whatever else causation in the case at hand consists in), would we still face a causing? Traditionally, too, analysis of this kind is highly intuition-driven. In my own view, intuitions are a decidedly unreliable guide to truths about possible worlds. Of course, we may use our modal intuitions as a cheap heuristic when the stakes are low—in the same way we rely on our geographical intuitions when it doesn't matter if we arrive late or get lost. But if the stakes are higher we had better seek evidence. And that means that empirical investigation should play a central role in attempts to devise theories of causality and explanation. Therefore I think that philosophical analysis cannot do very much for us—against Humphreys' hope to the contrary. I thus completely concur with Salmon when he says:

> we cannot get very far in attempting to understand scientific explanation if we try to articulate a universally applicable logic of scientific explanation. What constitutes an adequate explanation depends crucially […] on the kind of world in which we live; moreover, what constitutes an explanation may differ from one domain to another in the actual world.

This, I submit, is much more in line with Salmon's overall empiricist stance. Only our world can tell us what an adequate explanation or causal relation consists in. If that is highly domain-specific, an knowable only *a posteriori*, then so be it.

4 An Alternative Reading of Salmon's Theory

In conclusion I would like to suggest an alternative reading of Salmon's theory. Let us regard probability-raising and the transmission of conserved quantities not as ingredients in a conceptual analysis or other *theory* of causation but rather as indicators for whether or not a statistical relationship or process is causal. Then think of the relation between causation and its indicators in terms of the relation between a disease and its symptoms (the presence of a virus or bacteria in a patient's body I would count also as a symptom here). Tests are always tests of whether or not an indicator is present. The relationship between disease and indicators is at best one to many and but may even be many to many (probability-raising in a homogenous reference class may arise due to a causal connection but it may also be "brute").[13]

On this picture, a "theory" of causation may still be necessarily or contingently true ("necessary" understood in the Kripkean sense) but for sure it would be even less than domain specific, it would be *case* specific. This is because there is no reason to suppose that the indicators of causation correlate perfectly with our disciplinary boundaries: a variety of tests may work for economics, at least some of which are also applicable to medicine, which in turn develops its own tests but these may be applicable to certain cases in epidemiology some of whose tests may again work in economics.

Thus we surely give up on the initial motivation for Humphreys' project, *viz*. to gain a deeper understanding of the causal relation or explanation. But an empiricist should be cautious

[13] Nancy Cartwright tells me (personal communication) that she has used the disease analogy in an earlier paper of hers.

about seeking such a deeper understanding anyway. Once we know the various symptoms of, say, the flu, we know how to test for them, we know what brings it about and how to attempt to prevent and cure it, how much else is there to know about what the flu really *is*?

References (consolidated)

Cartwright, Nancy [1983], *How the Laws of Physics Lie*, Oxford: OUP.

Cartwright, Nancy [2003], "How Can We Know What Made the Ratman Sick? Singular Causes and Population Probabilities. An Essay in Honour of Adolf Grunbaum", *Causality: Metaphysics and Methods Technical Report* 08/03, CPNSS, LSE.

Dowe, Phil [1992a]: 'Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory', *Philosophy of Science* **59**, pp. 195-216.

Dowe, Phil [1992b]: 'Process Causality and Asymmetry', *Erkenntnis* **37**, pp. 179-196.

Friedman, Michael [2001]: *The Dynamics of Reason*. Stanford, California: CSLI Publications.

Hempel, Carl [1948]: 'Studies in the Logic of Explanation', *Philosophy of Science* **15**, pp. 135-175. Reprinted as Chapter 10 of Carl Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press, 1965.

Hitchcock, Christopher [1995]: 'Discussion: Salmon on Explanatory Relevance', *Philosophy of Science* **62**, pp. 304-320.

Hoover, Kevin [2001], *Causality in Macroeconomics*, Cambridge: CUP.

Humphreys, Paul [2000]: 'Review of *Causality and Explanation* by Wesley C. Salmon', *Journal of Philosophy* **97**, pp. 523-527.

Humphreys, Paul [2004]: *Extending Ourselves: Computational Science, Empiricism, and Scientific Method.* Oxford: Oxford University Press.

Lewis, David [1973]: 'Causation' *Journal of Philosophy* **70**, pp. 556-567.

Lewis, David [1986]: 'Postscripts to "Causation"', pp. 172-213 in David Lewis, *Philosophical Papers, Volume II*. Oxford: Oxford University Press.

Ryckman, Thomas [forthcoming]: *The Reign of Relativity.* Oxford: Oxford University Press.

Salmon, Wesley [1970]: 'Statistical Explanation', pp. 173-231 in *Nature and Function of Scientific Theories,* Robert G. Colodny (ed). Pittsburgh: University of Pittsburgh Press.

Salmon, Wesley [1985]: 'Scientific Explanation: Three Basic Conceptions', pp. 293-305 in *PSA 1984, Volume 2*, Peter Asquith and Philip Kitcher (eds). East Lansing, Michigan: Philosophy of Science Association. Reprinted as Chapter 20 in Salmon 1998.

Salmon, Wesley [1994]: 'Causality Without Counterfactuals', *Philosophy of Science* **61**, pp. 297-312. Reprinted as Chapter 16 in Salmon 1998.

Salmon, Wesley [1997]: 'Causality and Explanation: A Reply to Two Critics', *Philosophy of Science* **64**, pp. 461-477.

Salmon, Wesley [1998]: *Causality and Explanation*. Oxford: Oxford University Press.

Salmon, Wesley [1998a]: 'The Importance of Scientific Understanding', pp. 79-91 in Salmon 1998.

Schaffer, Jonathan [2000], "Trumping Preemption", *Journal of Philosophy* **97**, 165-81.

Sober, Elliott [2001], "Venetian Sea Levels, British Bread Prices, and the Principle of the Common Cause", *British Journal for the Philosophy of Science* **52**, 1-16.

Tooley, Michael [1990], "Causation: Reductionism versus Realism", *Philosophy and Phenomenological Research* **50** (Supplement), 215-36.