

The Causal and Unification Approaches to Explanation Unified—Causally

MICHAEL STREVENS
Stanford University

Abstract

Two major modern accounts of explanation are the causal and the unification accounts. My aim in this paper is to provide a kind of unification of the two, by using the central technical apparatus of the unification account to solve a central problem faced by the causal account, namely, the problem of determining which parts of a causal network are explanatorily relevant to the occurrence of an explanandum. The result is a causal account of explanation that has many of the advantages of the unification account.

1. Introduction

The two major modern approaches to scientific explanation are the causal and the unification approaches. Each has its advantages. I will propose a new account of explanation—the *kairctic account*—that takes something from both approaches. Despite its borrowing from the unification account, the kairctic account is not a hybrid account of explanation, but a fully causal account: the technical apparatus of the unification approach is appropriated not for its ability to capture the notion of unification, but rather, for its ability—quite unintended by its creators—to capture a notion crucial to the causal account, namely, the notion of *difference-making*.

In what follows, I introduce the causal and the unification approaches to explanation (section 2), I explain why the causal approach needs the notion of difference-making and why no account of difference-making to date is entirely satisfactory (sections 3 and 4), and I show that the technical apparatus of unification provides, after a few modifications, a surprisingly powerful criterion for difference-making (section 5). The paper's final section argues that the new account of causal explanation developed in this paper inherits many of the advantages of the unification account.

Throughout, I am concerned exclusively with the explanation of events. Explananda such as laws must wait until another time.

2. Two Approaches to Explanation

2.1. *The Unification Approach*

On the unification approach, an event is explained by deriving the occurrence of the event using a theory that unifies many diverse phenomena, and thereby showing that the event is part of a very general, perhaps utterly pervasive, pattern of events in the universe. In the best-developed unification account of explanation, due to Philip Kitcher, an event is explained by deducing it using the theory that unifies the phenomena better than any other (Kitcher 1981, 1989). This most unifying theory can have several parts—it might consist of a physics, a biology, a psychology, and so on—but the fewer its parts, the more unified, and hence the more unifying, is the theory.

There is more than one definition of unifying power in the literature; I will present a characterization of unifying power that captures what is common to the rival definitions, a characterization whose simplicity and inclusiveness make up for its lack of precision.

The unifying power of a theory increases in proportion to the following properties of the theory:

1. *Generality*, the number of actual phenomena that can be derived using the theory,
2. *Simplicity*, and
3. *Cohesion*. This third desideratum has been characterized in a number of ways. The aim of the desideratum is to discriminate against theories that, rather than picking out real patterns of phenomena, pick out mere unpatterned conjunctions of phenomena (or perhaps even worse, all possible phenomena). Kitcher's version of a cohesion principle is his *stringency* requirement. Morrison (2000) presents an account of cohesion (for non-explanatory purposes) on which cohesive theories use a single parameter or set of parameters in all their explanations.¹

The theory with the greatest unifying power, then, is the theory with the optimal combination of generality, simplicity, and cohesion.

2.2. *The Causal Approach*

On the causal approach, what explains an event is the event's causal history. The notion of a causal history is a vague one—removing this vagueness is a part of the motivation for the discussion that follows—but it includes, by anyone's lights, various events that bear causally on the occurrence of the explanandum and the causal laws in virtue of which they do so. (Unlike

some writers, I use the term *causal law* very liberally, to refer to any generalization describing a systematic causal dependency.)

The principal motivation for the causal approach is the recognition, stemming in part from certain famous counterexamples to Hempel's deductive-nomological account of explanation, that the direction of explanation runs parallel to the direction of causation. The height of a flagpole explains the length of its shadow, but the length of the shadow does not explain the flagpole's height. This explanatory asymmetry mirrors a causal asymmetry; proponents of the causal approach to explanation take this as reason to think that causation is the fundamental ingredient of explanation.

There is little agreement as to the technique that ought to be used to pick out the explanatory parts of the causal story from the tangled web of causal influence. If anything, the best known versions of the causal approach—in particular, Salmon (1984) and Railton (1981)—take a *laissez-faire* approach, on which it is pure pragmatics that determines which elements of the causal web are mentioned in response to a given explanatory request; see section 3.3.²

2.3. *Some Advantages of the Unification Approach*

Let me briefly describe a class of explanations that the unification approach is well suited, and the causal approach poorly suited, to account for. By the end of the paper (section 6), I will have claimed this advantage of the unification approach for the causal approach: by appropriating the technical apparatus of the unification approach, the causal approach will have acquired the ability to give a satisfactory account of these previously problematic explanations.

The kind of explanation I have in mind is that in which causal details are either omitted or distorted. The unificationist is able to handle such cases easily, by showing that the omission or distortion of the details enables a far greater degree of unification than would otherwise be possible. The causalist is embarrassed, because giving the correct causal details is, on the causal approach, just what explanation is supposed to be.

There are several different classes of examples in which details are omitted or distorted. I mention three.

First is a kind of case in which a mathematical fact about a physical configuration seems to be doing most or all of the explanatory work. Kitcher (1989, 426) provides a useful example: the explanation of a particular person's inability to untie a certain knot may lie in a topological fact about the knot, rather than in a detailed causal story about the attempt at disentanglement. Putnam's comments on the proper explanation of a square peg's not fitting in a round hole afford another example (Putnam 1975, 295–297).

Second is the case of equilibrium explanation, where the final state of some system is explained by showing that it is the system's equilibrium state.

Such explanations often make no reference to the initial conditions of the system, or to the particular trajectory taken from those initial conditions to the final state. Thus, they make no reference to the particular causal process that produced the state. Sober (1983) argues that equilibrium explanation is not causal at all. This is perhaps too extreme, but it is certainly hard to see how the causal approach will accommodate equilibrium explanation's omission of causal detail.

Third and finally, the causal approach has great difficulty in accounting for idealization in explanation. Consider, to take a simple but classic example, the use of ideal gas theory to explain Boyle's law (an explanation of a law, not an event, but never mind). The ideal gas theory deliberately misrepresents the causal process underlying Boyle's law in three important respects:

1. It ignores the long range attractive forces between molecules,
2. It represents molecules as not colliding (since they are infinitely small), when in fact they do collide, and
3. It is a classical theory, yet the underlying processes are quantum processes.

As hard as it may be for proponents of the causal approach to explain the omission of certain causal details, it seems harder still to explain these causal distortions. Yet they appear in all the standard textbook explanations, basic and advanced, of Boyle's law.

The unification approach seems well poised, as I earlier remarked, to make sense of these cases, since it does not aim for causal accuracy, but rather for a kind of generality that may well be enhanced by omitting or changing the details. The causal approach, by contrast, imposes no generality desideratum to counterbalance the drive for causal accuracy. Explanation is just telling the causal story; but how, then, can the neglect or the mutilation of the story serve an explanatory need?

3. The Problem of Explanatory Relevance

What explains an event, on the causal approach, is something about the causal process that produces the event. It seems that not every causal fact is explanatory, however. The gravitational forces exerted by distant stars have a causal influence, however slight, on the outcomes of causal processes here on Earth, but they are not cited in the explanations of those outcomes. Either they are counted as completely irrelevant, or at least, as far less relevant than whatever other factors are cited. This suggests that a complete causal account of explanation must include a criterion for assessing explanatory relevance.

3.1. *Oh, Those Russians*

To better appreciate the need for a criterion of explanatory relevance, consider the famous story of the death of Rasputin. Rasputin's assassins served him poisoned teacakes, which failed to kill him. They shot him twice; still he did not die. Finally, they tied his hands and feet and threw him through a hole in the ice into the river Neva. He drowned.

Let us take this story at face value. Somehow, the poison and the gunshot wounds were not enough to end Rasputin's life. It was only when he was tied up and thrown into a freezing river, a fate that no human could escape, that he died. The story is entertaining and macabre precisely because, of all the attempts to kill Rasputin, only the final, virtually infallible attempt succeeded. This observation can be couched in terms of explanation: the poisoning and the shooting do not explain Rasputin's death; what explains his death is his being bound and thrown in the river.

The challenge to the causal approach, in a case such as this, is to pick out, from the complex causal network culminating in Rasputin's death, just the throwing in the river as the proper explanation of death. The poisoning and the shooting must be ignored, or at the very least, relegated to a subsidiary role. Also to be ignored are the countless other incidents, properties and things that had some causal influence on the way the story unfolded: the length of Rasputin's beard, the day's pollen count, the gravitational influence of Mars, and so on.³

3.2. *Explanatory Relevance as Difference-Making*

It is my contention that the parts of the causal network that are explanatorily relevant to Rasputin's death are just those parts of the network that *made a difference* to whether or not Rasputin died. The gravitational influence of Mars of course made no difference to whether or not the death occurred; it is, in general, not the sort of thing that normally makes a difference to humans living or dying. Poisoning, by contrast, is the kind of event that often makes a difference to whether a person lives or dies. What gives Rasputin's death its tabloid quality, however, is that in this particular case, the poison made no difference to the death (or at least, that is how the story is told). What did make a difference is Rasputin's being tied up and thrown in the river; this, then, is the part of the causal network that truly explains the death.

To generalize from the case of Rasputin, I propose that the explanatorily relevant parts of any causal network are the elements that made a difference to whether or not the explanandum occurred. It is important to note the *whether or not*. To be explanatorily relevant, a causal factor must not merely make a difference to *how* the explanandum occurred; it must make a difference large enough to bear on whether or not it occurred at all. It follows that, on the difference-making account, explananda are conceived of as events with relatively coarse individuation conditions. In the Rasputin

case, the event is that of Rasputin dying (or perhaps better, dying within a given time frame). Had Rasputin not been shaking, due to the poison, as he drowned, he would still have died, thus the explanandum would still have occurred, though a much finer-grained event, individuated in part by the particular motions of Rasputin's body, would not have occurred. Poison thus made no difference to whether or not Rasputin died, though it did make a difference to whether or not he died in exactly the way he did.

How are the individuation conditions for the explanandum to be determined? They are specified, either explicitly or implicitly, in the explanatory request. I may ask why Rasputin died, or I may ask why Rasputin died in exactly the way he did. These questions ask for explanations of two different events, and will therefore elicit two different explanations, the latter containing far more causal detail than the former, since there will tend to be many more difference-makers for a fine-grained event than for a coarse-grained event.⁴

The idea that difference-making is a necessary condition for explanatory relevance is not a new one.⁵ In section 4, I look at some standard accounts of difference-making, and argue that they encounter difficulties in dealing with the Rasputin case. I then develop an account of difference-making that will turn out to bear a striking resemblance to the machinery of the unification account.

3.3. Other Approaches to Explanatory Relevance

Before I continue, I will mention, without arguing against, two alternatives to the difference-making approach to explanatory relevance.

A very minimalist approach to explanatory relevance holds that all causal influences on an event are relevant. On this view, Rasputin's being poisoned and even the gravitational influence of Mars are explanatorily relevant to his death. Railton (1981) comes close to advocating this position, and Salmon (1984) seems to sympathize.

The minimalist view must account for our explanatory neglect of the influence of Mars by appealing to some sort of pragmatic considerations; Mars, the story will go, though relevant in principle, is somehow never contextually salient. What the considerations might be, I do not know.

A quite different approach to explanatory relevance looks to a more exclusive conception of the causal relation, on which it is simply false that the influence of Mars is causally connected to Rasputin's death. The right sort of relation could be the relation asserted to hold by a causal claim of the form *C is a cause of E*: it seems correct to say that the gravitational influence of Mars was not, whereas the throwing in the river was, a cause of Rasputin's death. Lewis (1986a) makes a proposal along these lines.

The challenge, on the Lewis approach, is to give truth conditions for causal claims. In my view, every causal claim is at root a causal-explanatory claim: to say that *C* was a cause of *E* is to say that *C* is a part of a causal

explanation of *E*. If this is correct, then to base an account of explanatory relevance on an account of causal claims is to get things the wrong way around: it is the account of causal claims that ought to invoke a theory of explanatory relevance.

I can agree with Lewis, however, that there is a very close relation between causal claims and explanatory claims. This is borne out by the fact that the candidate accounts of explanatory relevance examined in section 4 are all familiar from the literature on causal claims. An ecumenist might even say that Lewis and I are pursuing the same goal under two rather different descriptions.

4. Approaches to Difference-Making

There are a number of senses in which one might say that a causal influence *C* on an event *E* *made a difference* to whether or not *E* occurred. Each of these senses constitutes a relation that might reasonably be called one of difference-making; the question is whether any of these difference-making relations is the relation underlying our judgments of explanatory relevance.

In what follows, I examine several kinds of difference-making relation, using the Rasputin case as a test. To pass the test, an account of difference-making must judge that neither the gravitational influence of Mars nor Rasputin's poisoning made a difference to Rasputin's death (or at least, that the difference made was negligible), while also judging that Rasputin's being bound and thrown in the river made a great difference.

I first aim to show that two competitors to my own account of difference-making fail the test, for well-known reasons. In such a small space, the arguments against these competitors cannot, of course, be decisive; I hope only to indicate certain obstacles to their success.

4.1. *The Probabilistic Relevance Test*

On a probabilistic relevance approach, *C* made a difference to the occurrence of *E* if it raised (or perhaps just changed) the probability of *E*.⁶ The probabilistic relevance criterion does a good job of eliminating as potential explainers minuscule causal influences such as the gravitational influence of Mars. Mars's exerting the influence that it did made no difference to the probability of Rasputin's death, hence, though it is a part of the causal network surrounding the death, it is not an explanatorily relevant part of the network, and so does not belong in the explanation of the death.

In the matter of Rasputin's poisoning, the probabilistic relevance criterion is less successful. Rasputin's being poisoned surely did raise the probability of his dying, thus, the probabilistic relevance criterion counts it as relevant when it is not. The problem, it seems, is this: probabilistic relevance is good at determining what kinds of factors are *typically* difference-makers,

but it cannot decide whether or not one of these factors actually made a difference in any particular case.

On behalf of the probabilistic relevance account, the following consideration might be advanced: since poison did not kill Rasputin, there must have been some fact or set of facts about him—his iron constitution, or what he had for breakfast, say—that neutralized the effect of the poison. But in the presence of such a neutralizer, the poison did not, after all, increase the probability of death, and so does not count as explanatorily relevant.

Observe, however, that we do not know for certain that there was such a neutralizer. Perhaps Rasputin was just lucky: the poison did raise the probability of his dying, but as it happened, he did not die. Since we do not know whether Rasputin survived due to a neutralizer or due to luck, we do not know whether the poison raised the probability of his death or not. Yet we are able to make a firm judgment that the poison was not explanatorily relevant to his death. The judgment must, then, be based on something other than the poison's failure to raise the probability of death.

4.2. *The Counterfactual Test*

On a counterfactual criterion for explanatory relevance, a causal influence C on an event E counts as having made a difference to whether or not E occurred just in case, had C not occurred, E would not have occurred.⁷ The counterfactual test is not a complete account of difference-making, as it provides no way to judge difference-making in the case of an event produced by an irreducibly probabilistic process. But I will ignore this problem; my own account of difference-making is limited in the same way.

The counterfactual test counts both the influence of Mars and the poisoning as explanatorily irrelevant to Rasputin's death. If Mars's gravitational influence had not been present, Rasputin would have died all the same. If Rasputin had not been poisoned, he would have died all the same. So far, so good.

But what about Rasputin's being bound and thrown into the river? If his conspirators had not killed him in this way, would he have lived? Arguably not: they would simply have killed him in some other way. Had he not been thrown in the river, then, he would still have died. Thus, his being thrown in the river is explanatorily irrelevant to his death, and so does not belong in an explanation of his death. Here the counterfactual criterion judges incorrectly.

This problem is, of course, already famous in the literature on the counterfactual analysis of causal claims, that is, claims of the form C is a cause of E , where it is known as the preemption problem (Lewis 1973). Various attempts have been made to solve the problem by amending the counterfactual criterion while retaining its counterfactual flavor (Lewis 1986c, 2000). Many doubt the success of these attempts.⁸ The ongoing difficulties of the counterfactual account in dealing with preemption seem reason enough to consider alternative approaches.

4.3. *A Modified Classical Approach to Difference-Making*

There are two ways I might now proceed. The first is to explain how the unification account of explanation can be adapted to give my favored account of difference-making, the kairetic account. The second is to motivate the kairetic account of difference-making independently of the unification account, and then to show that the unification account approximates the kairetic account. I choose the latter course. In this section, I take what might be called, in view of its provenance in the logical empiricist tradition, a classical approach to the problem of explanatory relevance, and I modify it in a non-classical way to provide what I argue is an account of difference-making superior to the probabilistic relevance and the counterfactual accounts. The modified classical account is then further amended in section 5 to produce the kairetic account itself.

What I call the classical account of difference-making is to be found in the literature on causal claims rather than in the literature on explanation, but because causal claims are in effect claims of explanatory relevance (see section 3.3), it is eminently applicable to the explanatory problem.⁹ John Mackie perhaps most elegantly encapsulated the idea in his notion of an INUS condition. Mackie (crediting the underlying idea to Mill) suggested that, when we say that *C* is a cause of *E*, we mean that *C* is an insufficient but non-redundant part of an unnecessary but sufficient condition for *E*. That is, *C* is an essential part of some set of actual conditions jointly sufficient for *E* (Mackie 1974).

Mackie had difficulty in characterizing the sense in which a set of conditions had to be sufficient for *E*. This was in part because he aimed to give a complete account of causality, not just an account of causal claims. Sufficiency, then, could not mean *causal* sufficiency. I, by contrast, mean exactly that: a set of conditions, including both events and laws, is sufficient, I stipulate, for the occurrence of an event *E*, just in case the conditions jointly entail the causal production of *E*. That is, the laws must entail that the events causally produce *E*.

The Mackie account, with this causal definition of sufficiency substituted for Mackie's definition, constitutes what I call the modified classical account of difference-making. The modified classical account assumes, I note, that the explanandum is deterministically produced; probabilistic explanation will not be considered in this paper.

According to the modified classical account, then, the difference-makers for an explanandum *E* may be determined by the following procedure:

1. Take the causal web in which *E* is embedded, and find a part of the web—a set of actual initial conditions and laws—that was in itself sufficient to causally produce *E*. This corresponds, loosely, to Mackie's unnecessary but sufficient condition for *E*. I will call the representation of the laws and conditions a *veridical deterministic causal model* for *E* (veridical because the conditions were actual causal influences on the production of *E*).

2. Remove from the causal model anything that is not necessary for the causal production of E . Assuming, for simplicity's sake, that the model consists of a set of propositions describing laws and initial conditions jointly entailing the occurrence of E ,¹⁰ this means removing everything from the model that does not play a role in the entailment of E . Note that to remove a causal factor is not to negate it. If an event C is removed from a model, the new model does not deny that C occurs; it is simply silent about C .

The initial conditions and laws that remain all made a difference to the occurrence of E . They correspond to Mackie's non-redundant parts of a sufficient condition for E ; taken separately, each corresponds to what Mackie calls an insufficient but non-redundant part of the sufficient condition.

I call this the *eliminative procedure* for determining difference-makers, and so for determining explanatory relevance. I call the end product of the eliminative procedure—a stripped down version of the model with which step (2) of the procedure began, containing only elements of the model that made a difference to the occurrence of the explanandum E —an *explanatory kernel* for E . The eliminative procedure takes any veridical, deterministic causal model for E as its input, then, and produces as its output an explanatory kernel for E .

Two remarks. First, the role played by logical entailment in the eliminative procedure as I conceive it is very different from the role it plays in the classical account of difference-making. In logical empiricist philosophy, logical relations such as entailment replace metaphysical relations thought to be questionable. In the classical account of difference-making, entailment stands in for what I have been calling the relation of causal production. Whatever scientifically important properties the metaphysically loaded notion of causal production was once supposed to have are asserted, by the classicists, also to be properties of the entailment relation.

On my modified classical account, by contrast, entailment is accorded a far humbler role, that of *representing* relations of causal production. The fact that certain true propositions concerning initial conditions and laws in a causal model entail an event E is used to represent the fact that, in the real world, the initial conditions causally produce E in virtue of the laws. Not every entailment represents a causal process; for example, C entails $C \vee D$, but this logical relation does not, if C and D are descriptions of events, represent a causal process, and so is accorded no significance at any stage of the eliminative procedure. The metaphysical facts about causal influence, then, constrain the significance of the logical facts about entailment. This feature of the modified classical account is put to use in section 4.4 to solve a well-known problem with the classical account.

My second remark clarifies the relationship between appearing in a kernel and being a difference-maker. In order for an event (or law) C to

qualify as a difference-maker for an event E , it is necessary and sufficient that C appear in a kernel for E . That is, it is necessary and sufficient that there exist some veridical causal model for E that contains C , and from which C is not removed by step (2) of the eliminative procedure. It is not required that C appear in every kernel for E . There may be veridical causal models for E which do not contain C (perhaps because the models take up the causal story after the occurrence of C), and there may be other veridical causal models for E that contain C , but from which C is removed by the eliminative procedure (perhaps because the models describe several different temporal stages of the causal process, any one of which is sufficient for the occurrence of E). None of this compromises C 's status as a difference-maker, provided that it appears in just one model for E from which it is not removed by the eliminative procedure.

This is all for the good, because no single kernel will contain all the difference-makers for a given event. At most, a kernel will contain all the difference-makers operating at a given time. It is an interesting question how the difference-makers for an event E appearing in different kernels can be put together to form a grand explanation for E containing all the factors explanatorily relevant to E 's occurrence, but the answer to this question will have to await another paper.

4.4. Successes of the Modified Classical Account

Let me now show how the notion of difference-making encapsulated in the eliminative procedure rules the influence of Mars and the poisoning irrelevant to Rasputin's death, and the influence of the throwing in the river relevant.

Suppose that we are given a veridical deterministic causal model for Rasputin's death that mentions, among other things, the influence of Mars. The elements in the model that entail that Rasputin dies have nothing to do with Mars, thus Mars can be removed from the model without affecting the model's entailment of death. Mars's influence, then, did not make a difference to the death.

This may suggest to some readers a well-known objection to the classical account of difference-making, which can be adapted to create difficulties, apparently, for the modified classical account.¹¹ Take some element in the Mars-mentioning model that does play a part in entailing death, say, the proposition D that Rasputin was thrown in the river. Replace D with the proposition $C \supset D$, where C describes the influence of Mars. This does not affect the model's entailment of the death, nor, apparently, the model's veridicality. Yet C cannot be removed from this new model without destroying the entailment of E . Thus C —the influence of Mars—is after all a difference-maker.

This is an effective objection to the classical account, but it is easily handled by the modified classical account, which imposes a non-classical constraint on explanatory kernels by requiring that they are veridical causal

models. The propositions that make up a veridical causal model for E must not only be true and entail E , recall; they must represent parts of the causal web in which E is embedded, and their entailment of E must correspond to a real causal process by which E is causally produced.

The model that serves as a counterexample to the classical account, which I will call M_{\dagger} , satisfies neither of these requirements. First, a proposition such as $C \supset D$ does not represent any part of the causal web. Rather, it is a disjunction involving two separate parts of the web.

Second—and this is the deeper point—the entailment of E involving $C \supset D$ does not correspond to a real causal process. Specifically, the entailment includes a step in which $C \supset D$ together with C entails D , but there is no causal process corresponding to this step. No physical theory attributes to C and $C \supset D$ the power to bring about D . Although M_{\dagger} entails the occurrence of E , it cannot be said to represent a causal process producing E .

For this and the previous reason, M_{\dagger} does not count as a causal model for E , and so cannot be used to obtain an explanatory kernel for E containing C . The influence of Mars, then, does not count as a difference-maker for the death.

What does the work here, note, is a constraint on causal models that is founded in physical facts about causal influence. I assume that these facts can be read off the true theory of everything. Some readers may worry that physics does not provide such information. This worry does not affect only my project, but any causal approach to explanation that claims to deal with the flagpole/shadow case and other such problems, since the causalist solution to these problems assumes a science rich enough in causal information to discriminate between those entailments that do and those that do not correspond to relations of causal influence. The proper place for the discussion of the worry, then, is in a discussion of the viability of the causal approach as a whole, not, as in the present paper, a discussion of competing versions of the causal approach. I therefore put the issue to one side.

Now consider Rasputin's poisoning. Suppose that we start with a veridical deterministic causal model for Rasputin's death that represents his being poisoned. Provided that the model also represents his being thrown in the river, the poisoning can be removed without affecting the entailment of death.

But what if Rasputin's being thrown in the river is not in the model? Can we construct a model that entails Rasputin's death, but does so in virtue of his poisoning? If so, then because poisoning could not be removed from such a model, it would count as a difference-maker. I must show, then, that a model that relies on poisoning to entail death is not possible, or more exactly, is either not veridical or does not genuinely entail death.

Consider the form of a model in which death is by poisoning. The model would cite Rasputin's having been poisoned in such and such a way, together with a deterministic causal law of the form: When a person is

poisoned in such and such a way, and conditions *Z* obtain, that person will die.¹² If the model is to entail death, then, it must also specify that conditions *Z* held. But conditions *Z* could not possibly have held in the actual death, because Rasputin *was* poisoned in such and such a way, yet he did not die. Thus a model that states that *Z* held is not veridical. To put it less abstractly: when Rasputin was poisoned, some condition must have held that prevented the poison from killing him. A model can only entail Rasputin's death by poisoning if it asserts that no such condition held. But then the model asserts a falsehood, and so fails to satisfy the requirement of veridicality.

Finally, consider Rasputin's being thrown in the river. This is a factor that we want to count as having made a difference to his death. Does the eliminative procedure oblige? Of course it does. Take a model that states only (a) that Rasputin was bound and thrown into the river, (b) that it is a law that people bound and thrown into the river in such and such conditions invariably die, and (c) that the conditions required by the law held. The model is veridical and entails Rasputin's death. His being thrown into the river cannot be removed from the model without invalidating the entailment. So it is a difference-maker.

What about the fact that so embarrasses the counterfactual test for difference-making, that if Rasputin's attackers had not drowned him they would certainly have killed him in some other way? On my account, the existence of a single veridical model for death from which Rasputin's being thrown in the river cannot be removed is enough to make his being thrown in the river a difference-maker. Other veridical models for death exist, citing the determination of Rasputin's attackers to kill him and so on, that do not need to mention his being thrown in the river in order to entail his death. But this does not subtract from the difference-making status of his being thrown in the river. The existence of these other models simply adds other factors to the list of difference-makers for Rasputin's death. His being thrown in the river certainly made a difference to his dying, but so, of course, did his attackers' murderous resolve.

The difference between the modified classical account and the counterfactual account that is decisive in this case, then, is that whereas on the counterfactual account, we look at the effect of removing a potential difference-maker *C* from a single veridical scenario entailing *E* (roughly, a model specifying the complete state of the world up until the time of the occurrence of *C*), on the modified classical account, we look at the effect of removing *C* from any number of different such scenarios, requiring only that *C* be essential in a single scenario.¹³ The modified classical account, then, imposes a weaker criterion for difference-making than the counterfactual account, which in the Rasputin case and other cases where there are "backup causes", is just what is needed.

4.5. *A Limitation of the Modified Classical Account*

I now turn to an important limitation of the modified classical account, which is a consequence of the fact that the eliminative procedure treats the

elements of a causal model as atomic: they can be removed, but not modified.

Suppose that I wish to explain the breaking of a window. As it happens, I threw a rather heavy brick at the window. Question: did the mass of the brick make a difference to the fact of the window's breaking? Answer: if you mean, did the *exact* mass of the brick make a difference, then no. The brick weighed 2 kg, but had it weighed 1 kg or 3 kg, the window would have broken just the same. However, if you mean did the *approximate* mass of the brick make a difference, then yes, it did. It is our explanatory practice, then, to say (a) the exact mass of the brick did not make a difference, but (b) the fact that the brick weighed between (say) 1 and 5 kg did make a difference. (There is an upper limit on the mass because, let's say, I could not have thrown a heavier brick far enough.)

The eliminative procedure seems unable to capture this aspect of our explanatory practice. The specification of the mass of the brick can be either left in the model, in which case the model specifies the exact mass of the brick, or taken out, in which case the model specifies no mass for the brick whatsoever. A model that specifies no mass for the brick cannot entail the window's breaking, since it does not rule out the possibility that the brick weighed 2 grams or 2 tonnes. Thus, the modified classical view declares the exact mass a difference-maker, contrary to our explanatory practice.

5. The Kairetic Account of Difference-Making

5.1. Abstraction

The problem of the brick's mass discussed at the end of the last section suggests that the modified classical account of difference-making ought to be amended so as to make possible an operation on a causal model more subtle than the wholesale removal of causal elements. I need an operation that transforms a model specifying an exact mass for a brick into a model specifying only an approximate mass for the brick: between 1 kg and 5 kg, say. The operation, it seems, consists in making certain descriptions in the model more vague. I call such an operation an *abstraction*.

To better characterize the notion of abstraction, I introduce a relation between models that I call *generation*. I will continue to assume that a causal model is a set of propositions making assertions about initial conditions and laws. Say that one model M generates another model M' just in case (a) all causal factors explicitly mentioned by M are also explicitly mentioned by M' , and (b) M' says at least as much as M , or a little more formally, every proposition in M is entailed by the propositions in M' . Intuitively, if M generates M' , then M' may be obtained by adding some additional causal details to M 's description of a causal process. Abstraction I define as the inverse of generation; that is, a model M is an abstraction of another model M' just in case M generates M' .

I now alter the eliminative procedure to allow not only removals, but also any other kind of abstraction (or almost any other kind; see section 5.3). The step to be amended is step (2) of the procedure defined in section 4.3, which used to read

2. Remove from the causal model anything you can without invalidating the model's entailment of the fact that *E* is causally produced.

and should now read:

- 2'. Make the causal model as abstract as you can without invalidating its entailment of the fact that *E* is causally produced.

The idea, then, is to continue making the model more and more abstract, until any further abstraction would result in a model that no longer guaranteed the causal production of *E*. What is left is an explanatory kernel, a model containing only causal factors and laws, or more exactly, only properties of causal factors and laws—such as the brick's mass's property of being between 1 kg and 5 kg—that made a difference to the occurrence of *E*.

5.2. *The Problem of Multiple Abstraction*

This new version of the eliminative procedure is not entirely satisfactory for two reasons, to be discussed in this and the next section.

Almost any causal model will contain more than one magnitude that might be specified vaguely. Even in the case of my brick-throwing, there is both the brick's mass and its velocity. The more abstract I make the specification of one of these magnitudes—the wider, that is, the range within which I specify the magnitude—the more my ability to abstract the other magnitude is constrained. This makes it impossible to say, as we would like, that the exact mass of the brick did not make a difference to the window's breaking, as I will now explain.

Assume that, in order to entail that the window breaks, a veridical causal model of the breaking must entail that the brick has a momentum of at least 3 kgms^{-1} and no more than 20 kgms^{-1} .¹⁴ Now suppose that a model *M* specifies that the brick has a mass between 1 kg and 5 kg. If *M* is to entail that the window breaks, it must limit the velocity of the brick to between 3 ms^{-1} and 4 ms^{-1} in order to guarantee that the brick's momentum falls within the required range. (The brick's momentum is the product of its mass and velocity.) Another veridical model, which stipulates a narrower range for the brick's mass—say, between 2 kg and 4 kg—can stipulate a wider range for the velocity, namely, between 1.5 ms^{-1} and 5 ms^{-1} . A veridical model that stipulates the widest possible range for the velocity, between 1.5 ms^{-1} and 10 ms^{-1} , limits the brick's mass to an exact value of 2 kg.

Each of these three models represents a possible end point of the eliminative procedure, in that, were the mass or velocity ranges of any of the

models widened, and so made more abstract, the model would no longer entail the window's breaking. It follows that the eliminative procedure does not give us a determinate conclusion about what aspects of my brick-throwing made a difference to the window's breaking. Because of the existence of an end point that specifies a precise value of 2 kg for the brick's mass, the eliminative procedure seems not even to rule out the possibility that the exact mass of the brick did, after all, make a difference to the window's breaking. Thus the eliminative procedure, as it stands, fails to capture an important element of our judgments about explanatory difference-making.

This is not just a problem about bricks, but about genuinely important scientific explanations: as I will argue in section 6, we want to say that the exact strength of the long range forces between molecules does not make a difference to the truth of Boyle's law, but the present problem, if unsolved, will prevent our doing so.

The difficulty, it seems, is that different sequences of abstractions lead to different end points. I suggest that the solution is to specify, not a process, but a uniquely desirable end point. That end point, I further suggest, for a given starting model M , is the most abstract veridical causal model that entails the explanandum and which is itself an abstraction of M .

The suggested criterion for difference-making, then, retains all the constraints from the eliminative procedure—generation of the starting model, veridicality, and entailment of the explanandum—but replaces a procedure, that of applying a series of abstractions subject to the constraints, with a goal, that of finding, in whatever way, the most abstract model satisfying the constraints. This constitutes a move from a procedural criterion for difference-making to an optimizing criterion for difference-making.

The new optimizing criterion requires a measure of abstraction that imposes enough of an ordering on causal models that there always exists a most abstract model satisfying the other constraints. For many purposes, the ordering inherent in the above definition of an abstraction, on which one model is more abstract than another if it generates the other, is adequate. The most abstract model satisfying the constraints, then, will be the unique model that generates all other models satisfying the constraints.

Let me apply the optimizing criterion to the brick-throwing case. There is one way to describe the mass m and the velocity v of the brick so that the resulting model is veridical, entails the window breaking, and generates the models described above and any other competitors:

$$3 \leq mv \leq 20.$$

A model incorporating this description, then, is the most abstract veridical, deterministic causal model for the window breaking, and the properties specified by the model are those that made a difference to the window's

breaking. It follows that, as desired, the exact mass of the brick did not make a difference. Nor did the fact that the brick's mass fell in some particular range. Rather, what made the difference is that the brick's *momentum* fell in a particular range.

With the move to optimization, the criterion for identifying a model containing only difference-makers begins, very slightly, to resemble unificationism's criterion for identifying the most unifying theory. The resemblance is closer than it may appear. To explain this claim, I will introduce a new method for quantifying abstractness. I will not defend the method in this paper; it is put forward only to illustrate the subtle connections between criteria for unification and criteria for difference-making.

On the new definition, the degree of abstractness of a model is proportional to the number of possible physical systems satisfying the model. More exactly, since a typical model specifies ranges for one or more real-valued parameters and so is satisfied by an infinite number of possible physical systems, abstractness is proportional to the standard measure (in the mathematical sense) of the set of possible physical systems satisfying the model. This captures the intuitive idea that a model is more abstract the more different ways there are to fill out the model's description of reality, hence, the less that the model constrains reality.¹⁵

There is obviously much more to be said about the new abstractness ordering, but for my purposes, it is sufficient to make two remarks. First, the new scheme for quantifying abstractness respects and reproduces the ordering imposed by the generation relation (in most cases at least; see note 15). If one model M generates another model M' , then all possible physical systems that satisfy M' also satisfy M . Thus M is, on the new scheme, at least as abstract as M' , as desired. For this reason, a model that is maximally abstract in the new sense will tend also to be maximally abstract in the old sense. Second, the new measure of abstractness is rather similar, though not identical, to the measure of generality used to quantify a theory's unifying power.

Let me conclude by summarizing the optimizing criterion for difference-making as it stands so far. Like the modified classical account, the optimizing criterion transforms any given veridical, deterministic causal model for an event E into the corresponding explanatory kernel for E , a model containing only difference-makers for E . The kernel corresponding to a model M for E is the causal model K that generates M , entails E , and is as general as possible.

The similarities and differences between the optimizing criterion and unificationism's criterion for a good explanation of E will be explored shortly; first, I have one important addition to make to the optimizing criterion.

5.3. *Cohesion*

Consider the following problem. Take two causal models that represent Rasputin's dying in two quite different ways, say, a death by drowning

due to his being thrown in a river, and a death by poisoning due to his being fed toxic teacakes. Suppose that one of these, the first, which I will call M , contains the real difference-makers for Rasputin's death; the other is therefore not veridical.

My concern is this. Put the two models together in a disjunction, so as to form a new model $M\ddagger$ that states that *either* Rasputin was thrown in a river etc., *or* he was fed poison teacakes, etc. $M\ddagger$ is veridical, since one of these chains of events did occur, as claimed, and it entails Rasputin's death, since both chains of events lead to death. And it is more abstract than M , because it is satisfied by all systems satisfying M (the throwings in the river) and more besides (the successful poisonings).

It seems, then, that the optimizing criterion for difference-making ought to prefer $M\ddagger$ to M . As a result, Rasputin's being thrown in the river will turn out not to be a difference-maker for his death; rather, the difference-maker is the more abstract event of his being either thrown in the river or poisoned.¹⁶ This is an unwanted conclusion. Even if the disjunction can be said in some extenuated sense to be a difference-maker, we want the disjunct—the throwing in the river—to be a difference-maker too.

The solution to the problem is, of course, to forbid abstractions that result in this kind of disjunction. There are many ways to enact the prohibition. Abstractions are already required not to introduce causal factors not in the original model (section 5.1); it is unclear, however, whether $M\ddagger$ ought to be regarded as introducing new causal factors or merely as describing old factors in an alarmingly vague way.

Another approach to solving the problem, which will cast further light on the appeal of unification, is motivated by the following observation. Although all systems satisfying the disjunctive model for Rasputin's death entail the death, they do so in two quite different ways. In some, the entailment goes by way of facts and generalizations about poisoning, in others, by way of facts and generalizations about drowning. As a consequence, the set of all causal factors mentioned in $M\ddagger$ and found in the systems that satisfy it can be divided into two non-overlapping (or only slightly overlapping) subsets, such that, for any particular system satisfying the model, the causal process leading to death will involve the factors in one of these subsets but not the factors in the other. For this reason, I say that the disjunctive model is not *cohesive*.

The cohesion of a model, then, is a measure of the degree to which the same kinds of difference-makers are active in every physical system satisfying the model. If a model is maximally cohesive, every causal element mentioned in the model plays a role in the causal production of the explanandum in every system satisfying the model. In a less cohesive model, some causal elements come into play in only some of the systems satisfying the model. Maximal cohesion is perhaps too much to ask of any reasonably abstract causal model, so I make cohesion a desideratum, not an absolute

requirement, of the difference-making criterion: we ought to try to maximize the cohesion of an explanatory kernel, but not at any cost. Cohesion may be traded off, then, for generality. A radically disjunctive model, I presume, scores so dismally for cohesion as to outweigh any likely benefit in generality.

As the reader might expect, I intend to draw a parallel between my cohesion desideratum and the unification account's desideratum of the same name. I do not claim that my cohesion desideratum does everything that the unification account's desideratum does. But it is a principle of broadly the same sort: insofar as the same kinds of causal elements entail the explanandum in every system satisfying a model, as my cohesion desideratum requests, the same causal pattern is instantiated in every system. Thus a model that is cohesive in my sense has as its instances causal processes all fitting the same underlying pattern. This is, of course, the goal (without the *causal*) of the unificationist version of the cohesion desideratum. Any respectable unificationist desideratum of cohesion will, for this reason, tend to do much of the same work as my cohesion desideratum.

5.4. *The Kairetic Account*

According to the revised and, for the purposes of this paper, final version of my kairetic account of difference-making, the explanatory kernel corresponding to a veridical causal model M for an event E is the causal model K that satisfies the following conditions:

1. K generates M , and
2. K entails E ,

and that best satisfies the following desiderata:

3. *Generality*: K is satisfied by as many physically possible systems as can be, and
4. *Cohesion*: As many of the causal elements mentioned in K as can be play a role in the causal production of E in any system satisfying K .

A causal factor is a difference-maker for an event E if it appears in any kernel for E .

There are three important differences between the kairetic criterion for determining an explanatory kernel and the unification criterion for a good explanation. First, the kairetic account insists, as unificationism does not, that the factors that explain an event play a role in the causal production of the event.

Second, the kairetic account does not impose a simplicity requirement.

Third, whereas unificationism's desideratum of generality requests that a model predict many *actual phenomena*, the kairetic generality desideratum

requests that a model be satisfied by many *possible systems*. The more important difference here is between the *actual* and the *possible*: on the kairetic account, a model can be very general even if it has only one actual instantiation, provided that there are many physically possible instantiations. The reason for the difference is that the two criteria are attempting to accomplish two very different tasks: the unification criterion, to find the model that entails the most actual phenomena; the kairetic criterion, to find the model that contains the smallest amount of causal detail.

All three of these differences point to genuine philosophical differences between the causal and the unification approaches to explanation. Yet, philosophy aside, the technical apparatus used by the kairetic account is surprisingly similar to the apparatus used by the unification approach. For this reason, the unification criterion is rather good at picking out difference-makers, and in particular, at identifying and throwing away irrelevant causal influences, such as the influence of Mars on Rasputin's death, which tend to compromise generality without increasing predictive power. Indeed, the unification account is in this respect considerably better at picking out difference-makers than the criteria underlying any causal account of explanation that I know of.

I suggest that the unification account is attractive to philosophers in large part because it is able to identify difference-makers, though that of course is not the intention of its proponents. The unification account unwittingly does something that the causal account should do, better than any extant version of the causal account can do it. I argue against the unification account and for the causal account, then, by praising the unification account—and showing that much of its goodness lies in its covert satisfaction of our causal needs.

6. Some Advantages of the Kairetic Account

The kairetic account enjoys many of the advantages attributed in section 2.3, in the discussion of knots, equilibria, and ideal gases, to the unification account. The reason, briefly, is that a difference-making theory of explanation is able to account, in a principled way, for the omission of causal details in a causal explanation: details are omitted because they made no difference to the occurrence of the explanandum.

Let me illustrate this claim by returning to section 2.3's examples. The first is Kitcher's knot. The proper explanation of a particular person's inability to untie an especially intractable knot, recall, may lie more in a topological fact about the knot than in the particular actions taken by the untier in the course of their effort. The kairetic account leaves out the specifics because they made no difference to the untier's difficulties: any of a large class of similar maneuvers would have been equally hopeless. What is important is that the untier attempted one or more of these doomed

maneuvers; which were chosen is not important, any more than it matters what knots, exactly, Rasputin's killers used to bind his hands and feet, provided that the knots were as secure as Kitcher's knot is convoluted.

The second example is equilibrium explanation, in which the initial conditions seem to play no role in accounting for a system's final state. According to the kairetic account, the initial conditions are omitted from the explanation because they made no difference to the final state: whatever the conditions were, the system was going to end up in the equilibrium state. The same applies to the omission of any description of the path taken from the initial state to the equilibrium state.

The third example, the use of ideal gas theory to explain Boyle's law, is rather more challenging, as it involves not the omission but the distortion of causal details. Ideal gas theory, recall, misrepresents causal details in at least three ways: it pretends that there are no long range forces between molecules, it pretends that there are no collisions between molecules, and it has a classical rather than a quantum aspect. How can the kairetic account make sense of the ideal gas model, which is clearly not veridical and so violates a basic condition for a good kairetic explanation?

A full answer to this question must await another paper, not least because the example involves the explanation of a law, but let me sketch the answer here. I begin with the first two misrepresentations. The ideal gas model is best understood, I hold, not as denying the existence of long range forces and collisions, but as ignoring them. Forces and collisions are, then, causal details that are omitted from a model because they make no difference to whether or not the explanandum, Boyle's law, holds.¹⁷ I suggest that the ideal gas model, properly understood, does not distort reality. What appears to be a distortion ought to be interpreted as a message about difference-making.

The same comment can be made, tentatively, about the third misrepresentation, the classical aspect of the ideal gas model. Whether the underlying physics of gas molecules is classical or quantum makes no difference to the form of the relationship between pressure and volume (though it makes a difference to other related phenomena, such as specific heats). A model containing only the difference-makers for Boyle's law must allow, then, for either a classical or a quantum physics of molecules. In practice, constructing such a model is not worth the trouble; we use instead a model that is gratifyingly simple and that in its willful misrepresentation of reality sends a strong message about what does and does not make a difference to the phenomenon it explains.

Inherent in this treatment of the omission and distortion of causal detail is a more general theory of the role, or at least of one of the roles, of idealization in science.¹⁸ Idealization is not just a crutch for scientists too weak-minded to comprehend reality in all its causal diversity. On the contrary, idealizations can represent a kind of positive knowledge about a

system, namely, the knowledge—critical to scientific understanding—of what aspects of a system’s workings do and do not make a difference to its overall behavior.

Notes

¹ Though there is some overlap between the requirements of simplicity and cohesion, the simplicity requirement cannot do all the work of the cohesion requirement. To see this, consider an example of Kitcher’s, the theory that whatever God wills, comes about. Any event can be derived from this theory, when combined with the appropriate initial condition (i.e., God willed the event). Furthermore, the theory is very simple. But there is no real pattern to the events derived from the theory, or at least, no pattern identified by the theory itself.

² Railton allows for non-causal explanation; I am referring to his account of those explanations that are causal.

³ The poisoning had an influence on how the causal story unfolded, but is it fair to say that it had a causal influence on the death itself? That very much depends on what notion of causal influence is intended. If causal relations are simply to be read off from physics, then the poisoning does have a causal influence on the death, insofar as the contents of the spatio-temporal region corresponding to the death are shaped in part by the poisoning: Rasputin may, for example, be shaking from the after-effects of the poisoning even as he drowns. I consider the possibility that a more exclusive notion of causal influence may solve the relevance problem in section 3.3.

⁴ My explananda are events in Kim’s sense, then, not in Davidson’s sense (Davidson 1969; Kim 1973). For an argument that scientific explananda are not Davidson events, see Hempel (1965, 421–23); also relevant is Lewis (1986c).

⁵ See, for example, Garfinkel (1981), Woodward (1984, 2003), and Hitchcock (1995).

⁶ Salmon (1970) is the best-known exponent of the view that explanatory relevance is probabilistic relevance.

⁷ Lewis (1986a) suggests what amounts to a counterfactual criterion for explanatory relevance; Woodward (2003) more explicitly does the same.

⁸ For my own doubts, see Strevens (2003). Woodward (2003) claims to solve the problem.

⁹ For a brief discussion of attempts to use something like the classical account of difference-making to solve relevance problems having to do explicitly with explanation, see Salmon (1990, 104). Especially notable among the late classicists is Fetzer (1981).

¹⁰ Not all actual scientific models have this structure, of course.

¹¹ The same kind of objection was leveled at Hempel’s deductive-nomological account of explanation; see Salmon (1990), section 1.1.

¹² I am, of course, oversimplifying here. There will not be a law of this form; rather, a generalization of this form will follow from many other, biologically more basic, laws.

¹³ There are, of course, a number of other differences between the two accounts.

¹⁴ Expository convenience rather than physical plausibility is my goal here.

¹⁵ The definition, I note in passing, is not useful for comparing models satisfied by sets of systems with infinite measure, as in the case where a real-valued parameter is unbounded in some direction. To deal with this problem simply, the new measure of abstraction may incorporate the old, using the generation relation to compare models satisfied by infinitely many systems. The details are unimportant, given my purposes here.

¹⁶ More exactly, the difference-maker is the disjunction of the facts characterizing the two causal chains in their entirety.

¹⁷ At least they make no difference to whether Boyle’s law holds approximately, which is presumably the explanandum here, since Boyle’s law does, in fact, hold only approximately.

Also, I should note that if intermolecular forces were great enough, Boyle's law would no longer hold even approximately; to say that the forces make no difference, then, is to say that they make no difference provided that they remain within certain realistic limits (compare the treatment of the parameters of the brick throw in section 5.2). There are other implicit limits to Boyle's law; it does not, for example, hold at very high pressures.

¹⁸ For a broadly similar view of idealization, see Batterman (2002).

References

- Batterman, R. W. (2002). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press, Oxford.
- Davidson, D. (1969). The individuation of events. In N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, pp. 216–34. D. Reidel, Dordrecht.
- Fetzer, J. (1981). *Scientific Knowledge*. D. Reidel, Dordrecht.
- Garfinkel, A. (1981). *Forms of Explanation*. Yale University Press, New Haven, CT.
- Hempel, C. (1965). Aspects of scientific explanation. In *Aspects of Scientific Explanation*, chap. 12. Free Press, New York.
- Hitchcock, C. R. (1995). Salmon on explanatory relevance. *Philosophy of Science* 62:304–20.
- Kim, J. (1973). Causation, nomic subsumption, and the concept of event. *Journal of Philosophy* 70:217–36.
- Kitcher, P. (1981). Explanatory unification. *Philosophy of Science* 48:507–31.
- . (1989). Explanatory unification and the causal structure of the world. In P. Kitcher and W. Salmon (eds.), *Scientific Explanation*, volume 13 of *Minnesota Studies in the Philosophy of Science*. University of Minnesota Press, Minneapolis.
- Lewis, D. (1973). Causation. *Journal of Philosophy* 70:556–67.
- . (1986a). Causal explanation. In Lewis (1986b).
- . (1986b). *Philosophical Papers*, volume 2. Oxford University Press, Oxford.
- . (1986c). Postscript to “Causation”. In Lewis (1986b).
- . (2000). Causation as influence. *Journal of Philosophy* 97:182–97.
- Mackie, J. (1974). *The Cement of the Universe*. Oxford University Press, Oxford.
- Morrison, M. (2000). *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge University Press, Cambridge.
- Putnam, H. (1975). Philosophy and our mental life. In *Mind, Language and Reality: Philosophical Papers*, volume 2. Cambridge University Press, Cambridge.
- Railton, P. (1981). Probability, explanation, and information. *Synthese* 48:233–56.
- Salmon, W. (1970). Statistical explanation. Reprinted in *Statistical Explanation and Statistical Relevance*. University of Pittsburgh Press, Pittsburgh.
- . (1984). *Explanation and the Causal Structure of the World*. Princeton University Press, Princeton, NJ.
- . (1990). *Four Decades of Scientific Explanation*. University of Minnesota Press, Minneapolis.
- Sober, E. (1983). Equilibrium explanation. *Philosophical Studies* 43:201–10.
- Strevens, M. (2003). Against Lewis's new theory of causation: A story with three morals. *Pacific Philosophical Quarterly* 84:398–412.
- Woodward, J. (1984). A theory of singular causal explanation. *Erkenntnis* 21:231–62.
- . (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.