INTERNATIONAL LIBRARY OF PHILOSOPHY

Cause and Chance

Causation in an indeterministic world

Edited by Phil Dowe
and Paul Noordhof

Cause and Chance

Causation is one of the oldest topics in philosophy, and has been a central problem for philosophers since David Hume. Most of the work done in this area has attempted to understand causation in deterministic worlds. But what about the unpredictable and chancy world we actually live in?

Cause and Chance: Causation in an Indeterministic World is a collection of specially written papers by world-class metaphysicians. Its focus is the problems facing the dominant 'reductionist' approach to causation: the attempt to cover all types of causation, deterministic and indeterministic, with one basic theory appealing to the notion of chance-raising.

This collection focuses on the two most substantial challenges the approach faces: the claim that chance-raising theories fail without an independent appeal to the notion of causal processes; and the claim that the standard practice of using counterfactuals to explain chance-raising doesn't work because counterfactuals themselves must be characterized in terms of causation.

Cause and Chance raises a number of further difficulties for reductive analyses and offers various ways of defending the chance-raising approach.

Contributors: Stephen Barker, Helen Beebee, Phil Dowe, Dorothy Edgington, Doug Ehring, Chris Hitchcock, Igal Kvart, Paul Noordhof, Murali Ramachandran, Michael Tooley.

Phil Dowe is Lecturer in Philosophy at the University of Queensland, and the author of *Physical Causation* (2000). **Paul Noordhof** is Reader in Philosophy at the University of Nottingham, and the author of *A Variety of Causes* (forthcoming).

International Library of Philosophy

Edited by José Luis Bermúdez, Tim Crane and Peter Sullivan

Recent titles in the ILP:

The Facts of Causation
D.H. Mellor

The Conceptual Roots of Mathematics
J.R. Lucas

Stream of Consciousness
Barry Dainton

Knowledge and Reference in Empirical Science
Jody Azzouni

Reason Without Freedom
David Owens

The Price of Doubt
N.M.L. Nathan

Matters of Mind
Scott Sturgeon

Logic, Form and Grammar
Peter Long

The Metaphysicians of Meaning
Gideon Makin

Logical Investigations, Vols I & II
Edmund Husserl

Truth Without Objectivity
Max Kölbel

Departing from Frege
Mark Sainsbury

The Importance of Being Understood
Adam Morton

Art and Morality
Edited by José Luis Bermúdez and Sebastian Gardner

Noble in Reason, Infinite in Faculty
A.W. Moore

Cause and Chance

Causation in an Indeterministic World

Edited by Phil Dowe
and Paul Noordhof

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2004
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2004.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

Cause and chance : causation in an indeterministic world / edited by
Phil Dowe and Paul Noordhof.
p. cm. – (International library of philosophy)
Includes bibliographical references and index.
1. Causation. I. Dowe, Phil. II. Noordhof, Paul, 1965– III. Series.

BD541.C194 2003
122–dc21

2003007275

ISBN 0-203-49466-0 Master e-book ISBN

ISBN 0-203-57269-6 (Adobe eReader Format)
ISBN 0–415–30098–3 (Print Edition)

Contents

Contributors

Stephen Barker is Lecturer of Philosophy at the University of Nottingham. He has been a research fellow at UNAM Mexico, Monash University and the University of Tasmania. He has written a book on speech-act semantics forthcoming with Oxford University Press, and is working on a book on counterfactuals and causation.

Helen Beebee is Senior Lecturer in Philosophy at the University of Manchester. She has published papers on causation, laws of nature, free will and epistemology in journals including *Mind*, *Journal of Philosophy*, *Noûs*, *Philosophical Quarterly* and *Analysis*.

Phil Dowe is Senior Lecturer in Philosophy at the University of Queensland. His book *Physical Causation* was published in 2000 by Cambridge University Press. His research interests, besides causation, include chance, identity, time and the interaction between science and religion.

Dorothy Edgington is Professor of Philosophy at Birkbeck College, University of London, and was Professor of Philosophy at Oxford University from 1996 to 2001. She is best known for her work on conditionals, including a long survey article, ‘On Conditionals’, in *Mind* 1995.

Douglas Ehring is Professor of Philosophy at Southern Methodist University in Dallas, Texas. His main area of specialization is metaphysics. He has published a book on causation with Oxford University Press, entitled *Causation and Persistence*, and numerous articles in journals including the *Journal of Philosophy*, *Noûs*, *Synthese*, *Philosophical Studies*, the *Australasian Journal of Philosophy* and *Analysis*.

Christopher Hitchcock is Professor of Philosophy at the California Institute of Technology. His research interests lie in the philosophy of science, especially in causation, explanation, probability and confirmation. He has published numerous articles in leading philosophical journals.

Igal Kvart is Professor of Philosophy at the Hebrew University, Jerusalem, Israel. He has published a series of articles on token probabilistic causation in leading philosophical journals. He was awarded the Johnsonian prize for his book *A Theory of Counterfactuals* (Hackett 1986). He has also published a series of articles on Reference, the de dicto–de re distinction, the attitudes (belief, knowledge and seeing) and their ascription, and mental causation. His home page is: http://socrates.huji.ac.il/Prof_Igal_Kvart.htm

Paul Noordhof is Reader in Philosophy at the University of Nottingham. He has written on causation in general, mental causation in particular, and issues connected with intentionality and consciousness. His book *A Variety of Causes* is forthcoming from Oxford University Press.

Murali Ramachandran is Senior Lecturer in Philosophy at the University of Sussex. His main areas of research are counterfactuals and causation, the theory of descriptions, rigid designation and contingent identity, and the analysis of knowledge.

Michael Tooley is the author of *Causation: A Realist Approach*, and of *Time, Tense, and Causation*. A former president of the Australasian Association of Philosophy, he is Professor of Philosophy at the University of Colorado at Boulder.

Chapter 1

Introduction

Phil Dowe and Paul Noordhof

The world most probably is *indeterministic*, meaning that there are particular events which lack a sufficient cause. Once we grant that there are such events, and that at least some of them are caused, we then require an account of causation that gives the conditions in which they are to count as caused. This is the problem of indeterministic causality. Providing for indeterministic causality has been a major motivation for the development of probabilistic accounts of causation.

A probabilistic account – essentially the idea that a cause raises the probability of its effect – is now commonplace in science and philosophy. It is taught as received knowledge in many fields. For example, in his medical textbook J. Mark Elwood offers this definition of cause: 'a factor is a cause if its operation increases the frequency of the event', and his caption describes this definition as 'The general definition of cause' (Elwood 1992: 6). Philosophers, on the other hand, have not been so sure. The papers in this volume contribute towards the proper articulation of the idea as well as, in some cases, subjecting it to sustained criticism. Below we briefly sketch some of the themes raised.

1 The characterization of chance-raising

Amongst philosophers who do agree that causes raise the chance of their effects, there has been disagreement over how this fundamental idea should be appropriately characterized. Some do so in terms of conditional probabilities (for example, see Igal Kvart, this volume); others do so in terms of subjunctive conditionals with chances of events figuring in the consequent of these conditionals (for example, see Paul Noordhof and Murali Ramachandran, this volume).

In philosophy, defining causes in terms of chance-raising was first made popular by Patrick Suppes' influential book *A Probabilistic Theory of Causality*, although both Reichenbach and Good had previously offered versions (Reichenbach 1956; Good 1961, 1962). Suppes defines prima facie causes, spurious causes, and genuine causes:

> Definition 1. An event B is a prima facie cause of an event A if and only if (i) B occurs earlier than A, (ii) the conditional probability of A occurring when B occurs is greater than the unconditional probability of A occurring.
>
> (Suppes 1984: 48)

For example, smoking (S) is a prima facie cause of lung cancer (C) because the conditional probability of getting lung cancer given that one smokes P(C|S) is greater than the unconditional probability of getting lung cancer P(C).

Not all prima facie cases turn out in the end to be genuine causes. Suppes therefore offers a definition of a spurious cause:

> Definition 2. An event B is a spurious cause of A if and only if B is a prima facie cause of A, and there is a partition of events earlier than B such that the conditional probability of A, given B and any element of the partition, is the same as the conditional probability of A, given just the element of the partition.
>
> (Suppes 1984: 50)

By 'a partition of events' Suppes means a way of dividing a kind of event into sub-kinds; for example, 'smoking' can be partitioned into 'occasional smoker/light smoker/heavy smoker' or smokers who drink and smokers who don't drink. For example, people are more likely to have a car accident if they are smoking than if they are not. Thus smoking is a prima facie cause of accidents. But if we introduce the partition high alcohol blood level/low alcohol blood level, we find that those who have high alcohol blood level and are smoking while they drive are just as likely to have an accident as those who have high alcohol blood level and are not smoking while they drive. It is also true that those who have low alcohol blood level and are smoking while they drive are just as likely to have an accident as those who have low alcohol blood level and are not smoking. So why the correlation between smoking and accidents? Just because it happens to be true, for whatever reason, that people who have been drinking are more likely to be smoking while they drive. In other words it's the drinking that causes accidents, not the smoking. So smoking turns out to be a spurious cause of accidents.

A prima facie cause that is not a spurious cause (that is there is no such partition to be found) Suppes defines as a genuine cause. For example, as far as we know there is no partition which would show that smoking is a spurious cause of lung cancer, so we regard it as a genuine cause. It is possible that such a partition will yet be found. Perhaps there is some condition that leads to lung cancer which makes people want to smoke.

Reference to events, causes and effects, may either be to types or tokens. Scientific laws, statistical and causal, concern the relationships between event types. However, there is also a need to consider the relationships between token events, for example in applied sciences such as medicine or engineering. A patient wants to know how *she* as an individual is going to be affected by such and such a treatment, and not just what happens in populations. In such cases there is an interest in the probability of effects as tokens. So an analysis of singular causation is called for. Suppes takes his account to apply to both tokens and types. Many of the contributors in this volume take themselves to be providing an account of token causation in the first instance. The exact relationship between accounts of token causation and type causation is controversial.

Philosophers working within the approach pioneered by Suppes have varied in the way they have dealt with the problem of other factors which reverse the probability relations. One approach is to take causation to be conditional on a certain set of background conditions:

$$P(E|C \ \& \ K) > P(E|{\sim}C \ \& \ K)$$

The problem remains to specify the right set of background conditions, or else leave causation as an essentially relative notion (Cartwright 1976). Others, for example Kvart, deal with this problem by conditionalizing on the entire history of the world up to the time of the cause (see Kvart, this volume).

As we noted, the other way to characterize chance-raising, more or less created single-handedly by the late David Lewis, is in terms of counterfactuals (Lewis 1986). First he defines a notion of probabilistic dependence.

> Event e_2 *probabilistically-depends* on a distinct event e_1 iff it is true that: if e_1 were to occur, the chance of e_2's occurring would be at least x, and if e_1 were not to occur, the chance of e_2's occurring would be at most y, where x is much greater than y.
>
> (Lewis 1986: 176–7)

This replaces notions of conditional probability with the idea of counterfactual chances: the chances an event has at different times and possible worlds. Given Lewis's story of what similarity in possible worlds amounts to, in normal situations this account deals with the problem of spurious factors by holding fixed the entire history up until just before the time of the cause, and seeing how the chance of the effect varies depending on whether or not the cause occurs. The phrases 'at least' and 'at most' have been introduced to try to accommodate Lewis's point that in the closest e_1-worlds – and also in the closest not-e_1-worlds – the chance of e_2 may fluctuate so that there is no precise chance that e_2 has. Lewis then defines causation in the following fashion.

> For any actual events e_1 and e_2, e_1 causes e_2 iff there are events $x_1, \ldots, x_n$ such that x_1 probabilistically depends upon $e_1, \ldots, e_2$ probabilistically depends on x_n.
>
> (Lewis 1986: 179)

The counterfactual theory handles many instances of indeterministic causality. The case where an insufficient cause is necessary in the circumstances exhibits both counterfactual and probabilistic dependence, while most cases involving an insufficient unnecessary cause exhibit probabilistic dependence. John's smoking causes his lung cancer, since in the closest possible worlds in which he doesn't smoke his chance of getting lung cancer is diminished.

Although some advance has been made on the proper characterization of chance-raising, there are significant difficulties facing the whole approach. It is to these that we now turn.

2 Chance-raising and causal processes

Few philosophers have been happy to accept that chance-raising is either necessary or sufficient for causation characterized in either way set out above (an exception is Hugh Mellor 1995). A number of familiar problem cases are responsible for this near consensus. One is preemption. Although it is just as much a problem for conditional probability-based accounts as for counterfactual accounts, it has been traditionally discussed in terms of the counterfactual theory. Another traditional feature of the debate, pioneered by the late David Lewis, is the appeal to neuron diagrams to provide a schematic model of the kind of case under discussion.

Suppose we have two possible chains of neurons both leading to the same event: *a–c–d–e*; *b–f–g–e*. Suppose that the first is a much more reliable process, that is the chance of *e* were *a* to occur is far greater than the chance of *e* were *b* to occur. Suppose also that *b*, on firing, may also inhibit *c*, thereby preventing the first process going through to completion. Take a particular case where *a* and *b* fire, *c* is inhibited, but, improbably, the second process goes through to completion, resulting in *e*. Figure 1.1 displays the scenario envisaged.

Intuitively, *b* causes *e*, but *e* does not probabilistically depend on *b*. Hence chance-raising is not sufficient for causation. Further, *a* does not cause *e* but *e* does probabilistically depend on *a* (Menzies 1989). Hence chance-raising is not necessary for causation.

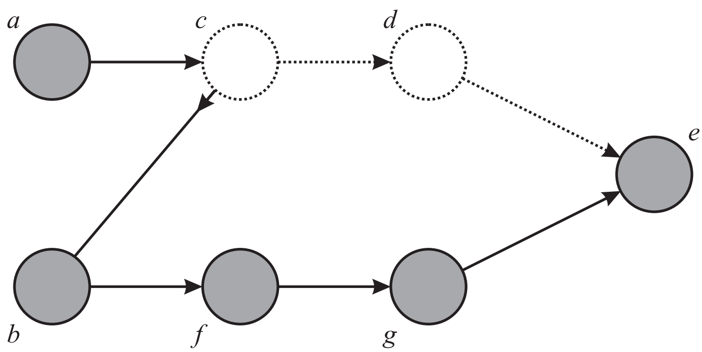

Figure 1.1 Early preemption

To deal with the first problem, Lewis allows for a chain of probabilistic dependence: *b* causes *e* because *e* probabilistically depends on *g*, *g* probabilistically depends on *f*, and *f* probabilistically depends on *b*. As Menzies shows, this does not solve the second problem, and so Menzies introduces an alternative theory. He suggests that 'e_1 causes e_2 only if there is a chain of unbroken causal processes running from e_1 to e_2' (Menzies 1989: 656).

The idea is that for any finite sequence of times $<t_1, \dots, t_n>$ between the time of e_1 and e_2, there is a sequence of actual events occurring at these times $<x_1, \dots, x_n>$ where x_1 is probabilistically dependent upon e_1, … e_2 is probabilistically dependent on x_n. Call this an unbroken causal process. A finite sequence of events $<a, b, c, \dots>$ is a chain of unbroken causal processes if and only if there is an unbroken causal process running from *a* to *b*, an unbroken causal process running from *b* to *c*, and so on. Talk of *chains* of unbroken causal processes is necessary to deal with the fact that e_1 can be a cause of e_2 even if there are some sequences of events between e_1 and e_2 which may not pairwise probabilistically depend upon each other. One example would be the finite sequence of events which just includes *b*'s firing and *e*'s firing in the original diagram (see Menzies 1989: 654–5; 1996: 93–4). Menzies's theory allows *b*'s firing to be a cause since unbroken causal processes can be patched together between *b*'s firing and *e*'s firing, whereas no chain of unbroken causal processes can be patched together between *a*'s firing and *e*'s firing.

Unfortunately, Menzies's account is inadequate – as he now recognizes. First, it rules out temporal action at a distance. It insists that there must be events at all the times between e_1 and e_2 for e_1 to cause e_2. Any theory which failed to rule this out a priori would have an advantage (Menzies 1996: 94).[1] Second, it cannot handle cases of probabilistic late preemption: that is cases in which the process preempted is preempted just by the occurrence of the effect. Consider the diagram in Figure 1.2.

As before, the *a*–*e* process is very reliable whereas the *b*–*e* process is unreliable.

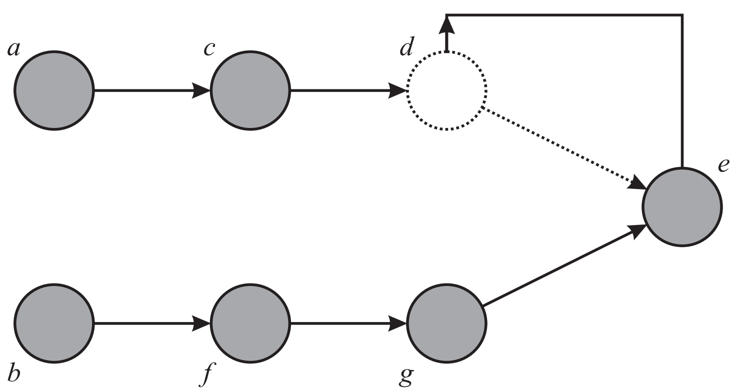

Figure 1.2 Late preemption

The crucial difference is that it is *e*'s firing which inhibits *d* from firing. If *e*'s firing had not occurred at the time it did as a result of the *b–e* process it would have occurred later – and hence after *d* firing – as a result of the *a*-chain. The problem is that there is a chain of pairwise probabilistically dependent events for all times between *a*'s firing and *e*'s firing but *a*'s firing is still not a cause of *e*'s firing because the *a*-chain had not completed: *e*'s firing occurred too early.

There have been other attempts to characterize when a causal chain is complete in counterfactual terms which avoids the problems identified with Menzies' proposal (see Ganeri, Noordhof and Ramachandran 1996; Ramachandran 1997; Ganeri, Noordhof and Ramachandran 1998; Noordhof 1999, Barker (this volume), Noordhof (this volume), and Ramachandran (this volume)). These attempts have focused on two features. First, if a causal process is incomplete then there may be a non-actual event, state or condition that would have completed the process. The accounts differ over whether it is appropriate to appeal to events, states or conditions broadly conceived (Barker says yes), or whether a more restrictive notion is necessary. Noordhof (this volume) claims that this is so and limits his appeal to events or states. Ramachandran also limits his appeal to events or states. It is partly for this reason that Barker insists that there *will* be a non-actual event, state or condition which would have completed the process whereas Ramachandran and Noordhof just suggest that there *may* have been. The latter two appeal to a second feature, which Barker eschews, to characterize incomplete processes. The question of whether or not putative causes affect the chance of the effect occurring at the time it did is assessed just before the time of occurrence of the effect. They suggest that this will not be the case in incomplete processes. This is a feature that Igal Kvart also emphasizes.

If these attempts by Noordhof and Ramachandran are to satisfy reductive ambitions, the counterfactuals to which they appeal should not have ineliminable reference to causation in their truth conditions. In her chapter in this volume, Dorothy Edgington argues that this requirement is not met. When we assess a counterfactual, we hold fixed facts which are causally independent from the truth or falsity of the antecedent right up to the time of occurrence of the consequent. Noordhof sketches one way in which Edgington's challenge may be avoided, but it is clear that the issue is a substantial one.

Igal Kvart has also argued previously that the semantics of counterfactuals contains ineliminable reference to causation (Kvart 1986). However, he does not suppose that this rules out a reductive account of causation but just the appeal of counterfactuals to supply it. In his contribution to the present volume, Kvart presents a sophisticated analysis of when a causal process is complete within the conditional probability framework. His key idea is that a complete causal process will be characterized by two elements. First, there will be events which, when they are taken into account, make a cause into a chance-raiser and no further actual events to reverse this fact (in his terminology, there will be stable increasers). Second, there will be no events which neutralize the chance-raising character of the cause unless they are, themselves, caused by the cause in question (in his terminology, stable screeners). Kvart explains how reference to causation in this characterization is not circular. He avoids one of the significant difficulties of conditional probability approaches – namely that the relevant conditional probabilities go undefined in deterministic worlds – by assuming that the world is indeterministic (see Lewis 1986). On the other hand, if Noordhof's attempt to defend the counterfactual position against this charge of circularity is successful, then a uniform account of deterministic and indeterministic causation would still be possible. That would be one consideration in favour of the counterfactual approach.

Some have argued that notions of conditional probability or counterfactual chances are insufficient for the proper characterization of causal processes and have advocated eschewing the reductive ambitions of such approaches. Instead, they urge that there should be unrepentant appeal to causal processes in a theory of causation. Wesley Salmon is the originator of this kind of position, with Phil Dowe, Doug Ehring and Michael Tooley its most prominent modern adherents. They advance their position significantly further in this volume.

They differ over their characterization of causal processes. Dowe offers an analysis of causal processes in the physical world in terms of conserved quantities. Ehring characterizes causal processes in terms of persisting tropes. Tooley takes causal processes to be theoretical entities defined in terms of being the unique satisfier of an open sentence stating various probability relations based on the fundamental idea that causes transmit their probabilities to their effects. In the present volume, Ehring advances his theory further by arguing that certain cases of causation can only be understood in terms of the idea of persisting tropes. In his contribution, Noordhof sketches a line of reply. In Chapter 6, Tooley argues that his theory can capture the only truth that there is to the link between chance-raising and causation, namely that the conditional chance of an event is higher, *given* that there is a law that would give it a positive probability, than its unconditional chance of occurring.

One question that some nonreductionists face concerns whether cases of prevention or hindering are cases of causation, since they do not seem to involve causal processes in the required sense. Phil Dowe has urged elsewhere that the answer is no (Dowe 2000b; 2001). Not all agree with this verdict (see Barker, this volume; Beebee, this volume; and Hitchcock, this volume).

3 Chance-lowering and causation

The proper characterization of complete causal processes is but one dimension of the question of whether causation should be linked to chance-raising. One way of putting the issue is to ask: however causal processes are to be characterized, is it true that there is a link between causation and chance-raising? Many of the contributors concede that causal processes don't always involve chance-raising but argue that they do in specific contexts. We might call this contingent chance-raising.

The case of preemption we discussed earlier illustrated the point. The unreliable process lowered the chance of the effect because it inhibited the more reliable process. Nevertheless, many share the intuition that the initial event in the unreliable process, in fact, caused the effect. In his paper for this volume, Dowe puts forward what he calls a path-specific approach. The basic idea is that, if we abstract a causal process away from the process which it hinders, then it will turn out to be chance-raising. By contrast, Hitchcock suggests that chance-raising is revealed if we hold fixed the events in the competing process. In articulating this idea, Hitchcock draws on Judea Pearl's influential recent work on the proper characterization of causal graphs (Pearl 2000). In their own contributions, Noordhof and Ramachandran argue that there will be some events such that, if we imagine them absent from the scene, chance-raising will be revealed. In certain circumstances, they take this chance-raising to indicate causation. Getting the proper characterization of contingent chance-raising right seems an important area for further study.

Scepticism about these approaches comes from two quarters. In a provocative article for this volume, Stephen Barker suggests that the kind of treatments sketched above to deal with preemption remove the motivation for appeal to chance-raising to characterize causation. Instead, one can appeal to counterfactual dependence. His own elegant proposal appeals to what he calls embedded counterfactual conditionals. We should consider counterfactuals in which counterfactuals which purportedly capture causal dependence occur in the consequent of an (embedded) counterfactual conditional. The antecedent of the whole counterfactual involves the supposition that certain aspects of the actual circumstances don't obtain. In brief, Barker argues that in abstracting away from circumstances, in order to reveal what others have thought to be chance-raising, you may in fact reduce the chance of an effect occurring in the absence of a cause to zero. Hence it will be possible to appeal to counterfactual dependence. Another who has defended this line, repudiated by him in the present volume, is Murali Ramachandran.[2] In his contribution, Noordhof questions whether this will always be the case. Barker's approach also differs from approaches such as that sketched out in Noordhof and Ramachandran's paper in that, as already noted, it appeals to embedded conditionals rather than conditionals with supplemented antecedents. As he notes at the end, there is no generally agreed semantics for such counterfactuals. As the semantics for embedded counterfactuals becomes more generally agreed, it will be interesting to see whether one approach is favoured over the other.

The other dimension of scepticism concerning the idea of causation being linked to contingent chance-raising is articulated in Beebee's challenging chapter. She suggests that none of these approaches will avoid the consequence that spraying defoliant on a weed is a cause of the weed's subsequent health. We will always be able to abstract away enough of the healthy plant processes so all that's left is the causal chain involving defoliation and health. In those circumstances, there will be contingent chance-raising. Hitchcock agrees with this verdict regarding Dowe's approach but rejects it for his own account. Noordhof also discusses the problem and sketches a line of reply. Beebee's own conclusion is that we should reject the idea of contingent chance-raising and just accept that all causation involves chance-raising. This involves the reclassification of some intuitive cases of causation as causal processes without causation but rather hindering (a distinctive kind of process). It seems clear from this discussion and from the brief earlier remarks about the status of prevention that the classification of *types* of causal processes and the characterization of their link to causation are matters of some importance.

Another matter which has received substantial attention recently is that of whether causation is transitive. Transitivity has not only been an article of faith but a means by which some of the problems regarding preemption seemed initially to be avoided. Even in his final work, Lewis insisted that, appearances to the contrary, causation is transitive. Many of the contributors are agreed that causation is not transitive but this causes potential difficulties in reductive accounts of indeterministic causal processes. Ramachandran brings this out very nicely in a range of striking, simple examples. For instance, in cases of mediate causation in which the chance of the effect is assessed just before the effect, the cause need not raise the chance of the effect because even if the cause had not occurred intermediate elements may have spontaneously occurred anyway. The fundamental contribution of his paper is to articulate one way in which those who are interested in providing a reductive account of causal processes may deal with the problem. He makes two moves. First, he writes into the antecedent of a counterfactual that putative effect should not have occurred before it actually does. He suggests that this will rule out spontaneous occurrences which obscure the probabilistic dependencies characteristic of causation. Second, although he appeals to chains of probabilistic dependencies he avoids claiming that causation is transitive by also requiring that causes are contingent chance-raisers where the chance of the effect is assessed just after the cause has occurred (in his terminology, they are early chance-raisers). Noordhof points to a problem with Ramachandran's approach and sketches an alternative way of dealing with the problem of indeterminism and transitivity in his own contribution. In it, he refines and defends the theory he set out in his *Mind* paper against some of the problems raised in this volume (Noordhof 1999).

The budget of problems identified by those who reject reductionism about causal processes, and indeed some friends of this project, is daunting. It would not be helpful to mention them all in the introduction. Table 1.1 summarizes where the main problems are discussed and who puts forward suggestions as to their solution within a chance-raising or counterfactual perspective.

As you will notice, some of the problems raised by Tooley do not receive further discussion in this volume. They will certainly require work in the future. Amongst the issues he raises in his rich paper are whether nonreductive theories of causation can capture causal asymmetries, whether causal relations supervene upon non-causal particular matters of act and laws, and what the metaphysics of objective chance is. In particular, he argues that if we take objective chances as fundamental, we will be committed to ruling out certain combinations of properties that otherwise seem compatible. Whether those who take causation to involve objective chance-raising must take these objective chances to be fundamental needs further discussion.

Table 1.1

Problems	*Discussed by*	*Solution proposed*
Early and late preemption	Dowe and Noordhof, Kvart, Ramachandran	Kvart, Ramachandran
Simple chance-lowering cases	Beebee, Dowe, Hitchcock, Noordhof, Tooley	Hitchcock, Noordhof
Trumping	Barker, Noordhof	Barker, Noordhof
Overlapping	Kvart, Noordhof	Kvart, Noordhof
Hastener–delayer asymmetry	Barker, Noordhof	Barker, Noordhof
Transitivity	Barker, Beebee, Noordhof, Ramachandran	Barker, Noordhof, Ramachandran
Causal asymmetry	Tooley	None!
Incomplete causal chains	Dowe and Noordhof, Noordhof, Ramachandran	Noordhof, Ramachandran
Spontaneous early occurrence of effect	Ramachandran	Ramachandran
Frustration	Barker	Barker
Overdetermination	Barker	Barker
Immediate action at a distance	Ehring, Noordhof	Noordhof
Objective chance	Tooley	None!

As an aid to placing the chapters in some kind of relation to each other in easily memorable form, we set out in Table 1.2 the conclusions on a range of key issues various chapters in this volume seem to favour. We hope you will get the impression that some significant issues have been identified and a measure of progress made.

Notes

1 For some considerations in favour of ruling out action at a distance, see Mellor (1995). For resistance, see Noordhof (1998b).

2 For criticisms of the attempt to provide a counterfactual approach which does not appeal to chance-raising see also David Lewis (1986) and Paul Noordhof (1998a). For a previous defence see Murali Ramachandran (1998). For criticisms in this volume, see Noordhof.

Table 1.2

Contributor	*Causes as chance-raising: straightforward, contingent or rejected*	*Chance-raising: conditional chance or counterfactual*	*Causal processes: reductive or nonreductive*
Stephen Barker	Rejected (but contingent counterfactual dependency)	Neutral	Nonreductive
Helen Beebee	Straightforward	Neutral	Neutral
Phil Dowe	Contingent	Neutral	Nonreductive: conserved quantity
Dorothy Edgington	Neutral	Neutral	Nonreductive
Doug Ehring	Rejected	Neutral	Nonreductive: persisting trope
Christopher Hitchcock	Contingent	Counterfactual	Neutral
Igal Kvart	Contingent	Conditional chance	Reductive
Paul Noordhof	Contingent	Counterfactual	Reductive
Murali Ramachandran	Contingent	Counterfactual	Reductive
Michael Tooley	Rejected	Neutral	Nonreductive: probability transmission

Chapter 2

Counterfactuals and the benefit of hindsight

Dorothy Edgington

I am driving to the airport to catch a nine o'clock flight to Paris. The car breaks down on the motorway. I sit there, gnashing my teeth, waiting for the breakdown service. Nine o'clock passes: I've missed my flight. More time passes. 'If I had caught the plane, I would have been half way to Paris by now', I say to the repairman who eventually shows up. 'Which flight were you on?' he asks. I tell him. 'Well you're wrong', he says. 'I was listening to the radio. It crashed. If you had caught that plane, you would be dead by now.'

With a bit more elaboration, which it will get in due course, this story is an example of a kind which creates a difficulty for all well-known theories of counterfactuals, and for a view I held. The problem is not new – it was mentioned in the 1970s – but it has received relatively little discussion until recently. Its force was impressed upon me by the work of Stephen Barker (1998, 1999). I had not entirely ignored it before, mentioning it in passing (Edgington 1995: 257–8), in criticism of David Lewis (1979). Yet later in the same article (I give details on p. 14–16), when saying what we aim at in counterfactual judgements – when such a judgement is objectively correct – I had forgotten about this sort of example, and so got things wrong. I shall try to rectify that. And I shall try to explain why we assess counterfactuals the way we do, in the context of an answer to the question: what do we need counterfactual judgements for? What purpose do they serve for us and why does it matter to get them right? What else goes wrong for us if we get counterfactuals wrong?

Before discussing the difficulty, I shall sketch the 'standard picture', common to various theories, and then my version of this standard picture. I shall only be concerned with counterfactuals whose antecedent and consequent are about particular states of affairs holding at particular times. Of course a theory needs to be more general than this, but that will not concern me here.

The standard picture

Goodman and Lewis

The problem of counterfactuals has always been: what are the rules of the game? (There is also the question: why play this game? And, if we answer that: which rules are appropriate to our purposes?) You suppose that something *A* had been true – something that is, often you know, actually false. You wonder whether, given that supposition, something else *C* would have been true. In trying to settle the matter, you need to rely on some actual facts, and let other actual facts go by the board with the supposition that *A*. What determines what you can hang on to, and what you must give up? Nelson Goodman (1955) gave us the form of a theory: '$A \Rightarrow C$' is true iff *C* is deducible from *A* together with the laws of nature together with facts which are cotenable with *A*. (I use '$\Rightarrow$' to symbolize the counterfactual conditional connective.) But he gave up when trying to specify which facts are cotenable with *A*. Something is not cotenable with *A* iff it would not be true if *A* were true. This is unacceptably circular.

At first, David Lewis's theory looked very different from the Goodman-style theories it succeeded. Call an *A*-world a world in which *A* is true. '$A \Rightarrow C$' is true iff *C* is true in all 'closest' *A*-worlds, that is all *A*-worlds which overall most resemble the actual world. Many readers of Lewis's book (1973a) assumed that ordinary common-or-garden standards of similarity were being invoked. Reviewers of the book, and others, pointed out that this doesn't work (see Bennett 1974; Fine 1975). By ordinary standards of similarity, the questions 'What would have happened if it had been the case that *A*?' and 'What is true in all *A*-worlds most similar to the actual world?' can get different answers. Any counterfactual of the form 'If *A*, then things would have been very different from the way they actually are' presents a difficulty. Kit Fine's example, discussed by Lewis (1979): if Nixon had pressed the button in 1974, there would have been a nuclear holocaust. But in the world most like the actual world in which Nixon pressed the button, nothing untoward happened. Lewis labels this the 'future similarity objection'. The objection is not answered just by discounting similarity after the consequent-time: its effect can arise from relying on similarity between antecedent-time and consequent-time. If Hitler had died in infancy, things would have been very different in the 1930s and 1940s. In the worlds in which Hitler died in infancy which most resemble the actual world up to the 1940s, however, some other child grows up to play a virtually identical Hitler-like role. In replying to these objections, Lewis (1979) was more explicit about which factors count towards closeness on the 'standard resolution' of the vagueness of the notion of similarity. His criteria are stated in more general terms to cover not only sequential counterfactuals about particular facts, but they have this consequence for the latter: the closest *A*-worlds are those with pasts identical to the actual world, up to shortly before the antecedent-time, when we need to deviate just enough to get the antecedent true. (Call the point of deviation the fork.) The closest *A*-worlds obey the laws of nature of the actual world, except insofar as we may need a small, inconspicuous deviation to get us to depart from the actual world at all. There is no deviation from the actual laws of nature after the fork. And that is all, or almost all. After we have deviated from perfect match, at the time of the fork, 'it is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly' (Lewis 1986: 48). (We shall return to this disjunction.) It is just (or almost just) the laws that we rely upon, after the fork.

This is what I shall call the standard picture. Note that the refinement of Lewis's theory brings it closer to Goodman's, with a strong hint about which facts are cotenable with the antecedent. Instead of saying that *C* is true in all *A*-worlds with the same particular facts up to the time of the fork and the same laws thereafter, we could say that *C* is deducible from *A* plus laws plus facts up to the time of the fork. Differences may show up if we consider a wider range of counterfactuals. Problems might arise about the short 'transition period' from the fork to *A*; but they usually don't. To a first approximation, they deliver the same picture. Others (such as Michael Slote 1978) have given Goodmanian theories somewhat along these lines.

Probabilistic version

The view of counterfactuals I held (inspired by Ernest Adams 1975) can be seen as a small modification of the standard picture described above. The truth conditions of the Lewis–Goodman type are, in my view, too strong. They make it too easy for a counterfactual to be plain false. Very many believable counterfactuals – possibly all or almost all the contingent counterfactuals we ever utter – could turn out to be downright false, on this version. I give three reasons:

1 Indeterminism

Suppose that all or many fundamental laws of nature are indeterministic: they may operate so as to make a certain outcome extremely probable given some conditions, but not certain. Then no or few ordinary propositions will be deducible from antecedent plus laws plus cotenable facts. *Mutatis mutandis*, no or few ordinary propositions will be true in *all* closest antecedent-worlds. To the extent that this is so, ordinary counterfactuals will be false, according to these truth conditions. If we believe that this is so, we should, on this theory, have no confidence at all in any counterfactual. I submit that, instead, if we believe that this is so, we should be less than completely certain, but are entitled still to be pretty confident – very close to certain, that if you had lit the gas, the water would have boiled, and so forth.

2 Determinism

Even if we do live in a deterministic world, we do not live in a crudely deterministic world. Our ordinary run-of-the-mill antecedents are not normally specific enough to be fed into deterministic laws. Even if coin-tossing is a deterministic process, no deterministic conclusion comes from the counterfactual supposition that you had tossed the coin, but only from a supposition of how *exactly down to the minutest detail* you tossed it. An example I have used to illustrate both these cases: a dog almost always, but not quite always, attacks and bites when strangers approach. We can detect no difference between the cases in which it does and those in which it doesn't. Assume either there's some indeterminism involved; or else, if there isn't, the outcome depends in some immensely subtle way on the manner of approach. I say 'I didn't approach, because I'm pretty sure that the dog would have bitten me if I had approached.' But on Goodman's theory, it is certainly false that if I had approached, I would have been bitten – either because of indeterminism, or, because the mere (coarse-grained) supposition that I approached, together with all cotenable facts and laws, does not entail that I was bitten. And what is certainly false is not something of which you should be close to certain.

The result is the same on Lewis's theory: assuming either indeterminism or fine-grained determinism, in almost all, but not quite all, close worlds in which I approach, I am bitten. That leaves the counterfactual clearly false.

3 Again assume determinism

The vocabulary in which the antecedent and consequent are couched may not be suitable for subsumption under the deterministic laws. This is particularly relevant to the countless counterfactuals we accept and assert about our own and others' mental lives. 'If I had received your invitation yesterday, I would have accepted.' Take the Davidsonian view. Even if determinism is true, these are not the categories which belong with the deterministic laws. Again, the assumption of the antecedent, together with other facts and laws, does not enable you to deduce the consequent. All such counterfactuals are false, on Goodman's and Lewis's theory. Whereas, it seems to me, our confidence (perhaps short of certainty) in counterfactuals such as these: 'If you had invited me yesterday, I would have accepted', 'If Mary had asked John to do the shopping, he would have done so', 'If Bill had been in London, he would have been in touch', does not depend upon our accepting that there are deterministic laws connecting consequent to antecedent and other relevant facts. Here is a perfectly ordinary use of a counterfactual: 'They're not at home; for the lights are off; and *if they had been at home, the lights would have been on*' (the example is used by Adams). You might be close to certain of the conditional, even if you are sure that their sitting in the dark is *not* inconsistent with the laws and cotenable facts. To repeat: on Goodman's theory, if you are sure that the consequent isn't entailed by laws and so on, you should be sure that the counterfactual is false.

I am *not* recommending that we say instead that a counterfactual is true iff the consequent is very probable given the antecedent, laws and cotenable facts. That won't work. Suppose we did say that. Suppose I know that it is indeed very probable that C would have been true if A had been true – say 95% probable; so I should be certain that it is true. So I should be certain that if A, C. But I'm not. I'm only close to certain. Suppose I know that it is not very probable that if A, C – it's around 50–50. Then I should be certain that it is not true. So I should have zero confidence that if A, C. But I don't: I think it is about 50–50 that C would have happened if A had. I'm suggesting instead that we simply stick with the appropriate conditional probability – the conditional probability of C given *A at the time of the fork*, as a measure of the acceptability of the counterfactual. You ask: how likely *was* it, *then*, that C would have happened if A had? One way of looking at it: consider all the Lewisian closest A-worlds. Suppose for simplicity that you have divided them into a finite number of equi-probable clumps in a suitable way. Then the question is, in what proportion of the clumps is C true? Whereas for Lewis, unless C is true at all the clumps, the counterfactual is plain false.

This view also fits with my view of indicative conditionals, and in particular vindicates a nice relation between typical forward-looking 'will'-conditionals and counterfactual 'would'-conditionals. We believe an indicative conditional to the extent that we think the consequent is probable on the supposition of the antecedent. For many forward-looking 'will'-conditionals, there is an objectively correct opinion to have: the objective chance of C given A. A boring and easy example: you are to pick a ball at random from a bag in which 90% of the red balls have black spots. What should you think about the conditional 'If I pick a red ball it will have a black spot'? You should be 90% confident that if you pick a red ball, it will have a black spot. That is the right, unimprovable opinion, at least before you pick. Suppose you do pick a red ball. Then this conditional probability will change – collapse – to 1 or 0. Suppose you don't pick a red ball. Then it doesn't collapse. It's 90% likely that if you *had* picked a red ball, it would have had a black spot. And there it remains, unalterable forevermore (or so I thought). Even God can't better that judgement.

In central cases, 'would'-conditionals and 'will'-conditionals differ merely in a temporal way: the same conditional thought can be expressed now with a 'will', later with a 'would'. I say 'Don't go in there; if you go in you will be hurt.' You look sceptical but stay outside, and there is a loud bang as the ceiling collapses. 'You see', I say, 'I was right: if you had gone in, you would have been hurt. *I told you so*.' Or, if there is no loud bang and the ceiling doesn't collapse, 'I was wrong; I thought the ceiling was about to collapse; I thought you would have been hurt if you had gone in.' 'If they're here by eight, we'll eat at nine' is rephrased hungrily at ten, 'If they had been here by eight, we would have eaten at nine.' I change my travel plans on being told, 'If you travel on Friday, it will cost you £20 extra.' I discover I was misinformed: if I had travelled on Friday, it would not have cost me extra. And so on. Your present 'would haves' agree with your present opinion about the acceptability of the corresponding earlier 'will'.

The above is the picture I presented in my 1995 paper, Sections 8 and 10. Section 8 concerned indicative conditionals, and argued that, despite lack of truth conditions, for many forward-looking indicatives, there is something objective to aim at: the objective chance of C on the supposition that A. If A turns out to be true, this chance collapses to 1 or 0, depending on whether C is true or false. If A turns out to be false, the objectively correct value to be assigned to the counterfactual, 'If it had been the case that A, it would have been the case that C' is the conditional chance of C on the supposition that A, at the time of the fork, just before it turned out that $\neg A$. Section 10 developed this theme for counterfactuals.

The problem

Return, at last, to the plane crash. Stipulate that a chance event, not predictable in advance, brought down the plane. Everyone aboard was killed. Indeed, there was no chance, after the crash, that anyone on board would survive. At the time of take-off, this plane was not relevantly different, with respect to safety, from any other normal plane: there was an extremely small but non-zero chance that some such accident would occur – due to freak weather conditions, or freak electrical or mechanical faults (or combinations thereof), or a freak heart attack or attacks on the part of those in control.

Is the repairman's remark correct? Well, perhaps not if, for example, some subtle feature of the distribution of weight in the plane played some causal role in the antecedents of the crash – a feature which might well have been different, had I been on board. But if, as is more likely, my absence from the plane had no effect on the aetiology of the crash, it is surely correct.

The first mention of an example like this in print is in a footnote at the end of a paper by Slote (1978), and is attributed to Sydney Morgenbesser. It is simply 'If I had bet on heads, I would have won', said of a presumed indeterministic coin-toss which landed heads. Similarly, any week after a lottery draw, I'm right, it seems, to say 'If I had chosen numbers 45 67 … I would have won'.

Slote says in this footnote:

> I know of no theory of counterfactuals which can adequately explain why such a statement seems natural and correct. But perhaps it simply *isn't* correct, and the correct retort is 'no, you're wrong; if I had bet (heads), the coin might have come up differently, and (so) I might have lost – assuming the coin was random'.
>
> (Slote 1978: 27)

This, I think, is wishful thinking[1] (wishful philosophical thinking, that is: the example refutes the thesis of Slote's paper). Consider: you are watching a lottery draw on television and to your dismay your arch business rival wins a prize – not a big enough prize for him to abandon his business, but big enough for him to put you out of yours. If Slote's suggested 'retort' were correct, so would this be: you say to yourself, 'If I had scratched my nose a minute ago, he very probably would have lost. What a pity I didn't scratch my nose!'[2]

We can do better than just to appeal to intuitions. In fact, the appeal to intuitions is compelling, I think. But it leaves hanging the question of *why* our counterfactual thought-experiments are conducted in this manner. In the final section of the paper, I try to show how the intuitive response to these examples is the one that fits the use we make of these judgements.

(Note: you have to countenance the possibility of indeterminism for these examples to be a problem for the standard view. It seems to me (and to Lewis) that a decent theory of counterfactuals should cater for that possibility. But for someone who thinks, on something like a priori grounds, that determinism must be true, the standard view is not in trouble: sufficient causes of the plane crash were there back before the fork.)

A few more remarks about the plane crash. First, it was not essential to the story that the crash was such that those on board had a 100% chance of being killed. Perhaps there were a few survivors. Perhaps there was about a 90% chance of being killed if on board. Even so, I will think, 'It's very likely that I would have been killed if I had been on board' – unless I can tell a special story about my abnormal powers of survival.

Second, as mentioned above, there can be mixed cases where there is some chance that my presence on the plane would have altered conditions in a way to prevent the crash, and some chance that it would not have interfered with the crash. For a purer example, consider a coin toss.

- Case 1: I decline to bet. It lands heads. Assume no causal interference. If I had bet on heads I would have won.
- Case 2: there is a cheat around. He has a little gizmo in the palm of his hand. When someone bets heads, he presses it, and it sends out a magnetic pulse or whatever, which prevents the coin landing heads. In this case, if I had bet on heads, the coin would not have landed heads.
- Case 3: we have a more sophisticated cheat. After all, suspicion would be aroused if coins *never* landed heads when people bet on heads. He does not trust himself to randomize. The device does it for him. He always presses it when someone bets heads. There's a 90% chance that it does nothing, and a 10% chance that it prevents the coin from landing heads. I didn't bet. The coin lands heads. If I had bet on heads, it's 90% likely that I would have won; for there was a 90% chance of no causal interference, and a 10% chance that the coin would have been prevented from landing heads. Note that here too, the way the coin *actually landed* carries weight in assessing the counterfactual. (Examples like this are discussed in Barker 1999.)

Finally, let me stress again the crucial role of *causal* independence. As a fantasy, imagine that the crash has this genesis: the devil spins a spinner, which has a one-in-a-million chance of landing in the space designated 'crash'. It does land there. There is a crash. If I had caught the plane I would be dead. Now suppose that the devil has two identical spinners, and some rule for deciding which to use which has the consequence that he will spin one if I am on the plane, the other if I am not. Although the chances are initially just the same, in this case, if I had caught the plane, very probably it would not have crashed!

The problem for the standard view, in any version, is, of course, that actual facts after the time of the fork can be crucially important to the assessment of counterfactuals.

The problem for Lewis

'It is of little or no importance to secure approximate similarity of particular fact [between worlds], even in matters which concern us greatly' says Lewis (1986: 48). That is, after the fork when we no longer have perfect match, similarity of particular fact is of little or no importance. The nearest he gets to addressing our problem is in the parenthetical remark which follows: 'It is a good question whether approximate similarities of particular fact should have little weight or none. Different cases come out differently, and I would like to know why. Tichy (1976) and Jackson (1977) give cases which appear to come out right … only if approximate similarities count for nothing; but Morgenbesser … has given a case which appears to go the other way.' That is all he says and he has never, as far as I know, returned to the problem. The Tichy example is this: when Fred goes out and it's raining, he always takes his hat. When he goes out and it's not raining, it's a random 50–50 whether he takes his hat. On this occasion, it's raining and he takes his hat. Consider 'If it had not been raining, he would have taken his hat'. The fine-weather world in which he does take his hat resembles the actual world more than the fine-weather world in which he does not take his hat does. But this, Lewis rightly wants to say, counts for nothing: the counterfactual is not clearly true. (For Lewis, it's clearly false; for me, it's 50–50.) This suggests the no-weight picture has to be the right one.

Many examples go the other way: if I had bet on heads, I would have won; if I had bought these shares a year ago, I would be rich; if I had left five minutes earlier, I would have avoided the accident; if I had got up five minutes earlier, the result of the Australian General Election would have been just the same.

I pick a coin from a bowl of coins, toss it, and it lands heads. It would be wrong to claim that if I had picked a different coin, it too would have landed heads. But it would be absurd to deny that if Frank in Australia had scratched his nose a moment or two earlier, the coin I picked, and tossed, which actually landed heads, would still have landed heads. The difficulty for Lewis is distinguishing these cases.

Another such pair. Guerrilla warfare in an imaginary country. The guerrilla leader is hiding in a certain village. Government troops have a range of missiles aimed at the village. These devices are indeterministic, and each has a chance of, say, 90% of firing when activated. News having arrived of the need to deploy troops elsewhere, only one missile is to be set off. The General chooses a missile, which is activated. It fizzles out. No harm is done.

'We were lucky', says a potential victim later. 'Had the General chosen a different missile, we might well be dead.' His companion, versed in Lewis's early work on counterfactuals, taking 'similarity' in an intuitive way, demurs. 'We were lucky that *it* didn't go off', he says, 'but your relief is misplaced. In the world most like the actual world in which he chose a different missile, it fizzled out too, right? That is to say, if he had chosen a different missile, it too would have fizzled out.' The silliness of this suggests that Lewis should say 'approximate similarity counts for nothing'.

Version two of the story: two inhabitants of the village are delayed on their way home because they notice a sheep caught in a cactus, and it takes them a while to free it. The scenario is as before, but let me lower the chance each missile has of firing, to about 25%. This time, the missile does fire. They hear it in the distance. When they get back, they meet havoc and destruction. 'If we hadn't noticed the sheep, we would probably be dead now', says one. His companion, versed in Lewis's later work and the reasons for saying 'approximate similarity counts for nothing' demurs: 'Consider the possible world which deviated from the actual one at the time we noticed the sheep: the missile (or its counterpart), in that world, had only a 25% chance of going off. So if we hadn't noticed the sheep, it's 75% likely that the disaster would not have occurred. What a pity we noticed the sheep! If we hadn't, probably, all would be well.'

I don't see how Lewis can handle these examples without appealing to the notion of causal independence. Whether Fred wears his hat is not causally independent of the weather. Picking another coin or missile begins a different causal process. But the outcome for this coin or missile that was picked is causally independent of someone scratching his nose in Australia, or the antics of the sheep. As Lewis wants to explain causal dependence and independence in terms of counterfactuals, this is a problem for him.

It might be thought that the standard picture gives the conditions for causation, even if it doesn't always agree with our counterfactual judgements. But the problem cases seem to show that the standard picture gives the wrong conditions for causation – at least when we allow causation to be indeterministic, as Lewis does, and as do all who pursue this approach to causation. Lewis's account went like this: *c* directly causes *e* iff *c* occurs, *e* occurs and the actual chance, immediately after *c* occurs, that *e* will occur, is significantly higher than the chance of *e* occurring in the absence of *c*. Suppose I'm facing a machine which emits particles. I snap my fingers. Immediately afterwards, the chance that a particle is emitted reaches 100%. Are these events causally related? We have to assess the counterfactual 'If I had not snapped my fingers, the chance would have been much less than 100% that a particle be emitted'.

To assess this according to the standard picture, we go back to the time of the fork, shortly before I snapped my fingers; we appeal to the actual laws of nature but not to particular facts thereafter; and we ask what is the chance, at a time just after the antecedent time, that a particle be emitted. It might be very low. So the standard picture delivers the wrong answer. (The example is Barker's, this volume.) Of course we want to say: the particle would have been emitted even if I had not snapped my fingers. But that *rests on* a judgement of causal independence.

(In addition to the point about the need to appeal to causation, I suggest that Lewis's modal realism makes it hard for him to put weight on the distinction that matters here. A die is tossed, and lands six. We can't infer that if another die had been tossed instead, it would have landed six. But we can infer that if Frank on the other side of the world had scratched his nose, this die would still have landed six. But for Lewis, this latter question is a question about whether, in all sufficiently close possible worlds in which counterpart Franks scratch their noses, counterpart dice in those worlds land six. It is hard to see what would *ground* the right answer, when the question is put in Lewis's terms.)

Handling the problem

If we give up on the idea of explaining causation in terms of counterfactuals (or never had that idea in the first place), it is not too hard to see how the standard picture needs to be amended to handle these examples. When we assess a counterfactual, we may need to take into account the way the world actually rolls on, after the fork, in ways which are causally independent of our antecedent. A Lewis-style account would go thus: consider those A-worlds which (a) depart from the actual world shortly before the time of $\neg A$, at an inconspicuous fork; (b) thereafter obey the actual laws of nature; and (c) share with the actual world subsequent particular facts which are causally independent of $\neg A$, up to the time of the consequent. A counterfactual $A \Rightarrow C$ is true iff C is true at all such worlds. A Goodman-style story will say: the cotenable facts are (a) those up to shortly before the time of $\neg A$; (b) the laws of nature; (c) any subsequent fact, up to the time of the consequent, which is causally independent of $\neg A$ (in other words whose causal history does not go through $\neg A$). $A \Rightarrow C$ is true iff C is entailed by A and the cotenable facts. I would subject both to probabilistic amendment, as before. Instead of Lewis's truth condition I would say (very crudely and roughly) that $A \Rightarrow C$ is probable to the extent that C is true in most of those A-worlds. Instead of Goodman's, $A \Rightarrow C$ is probable to the extent that the chance is high, at the time of the fork, of C given A and the cotenable facts. The objectively correct value to assign to such a counterfactual is not (or not always) the conditional chance of C given A at the time of the fork; but the conditional chance, at that time, of C given A & S where S is a conjunction of those facts concerning the time between antecedent and consequent which are (a) causally independent of the antecedent, and (b) affect the chance of the consequent.

At first sight this is a rather strange probabilistic animal, but it is a bona fide conditional probability. Think of it this way. Go to a time just before $\neg A$, and consider the chance then of C given A. There is a future of branching paths, some of which lead to C and some of which lead to $\neg C$, and with other possible intervening events along the various paths. For any actual fact S causally independent of A which affects the chance of C, cross out the $\neg S$ paths and recalculate the chance of C given A on that basis.

Figure 2.1 illustrates this for the case of the crash. At the time of the 'fork' when there is still some chance that I catch the plane, the chance of a crash is very low, and hence the chance that I will die if I catch the plane is very low. I miss the plane. There is a crash, which is causally independent of my presence or absence from the plane. In assessing the counterfactual 'If I had caught the plane, I would have died', we now eliminate the 'no crash' paths. The relevant chance is the chance, at the time of the fork, that I die given that I catch the plane, and given any subsequent relevant causally independent facts – that is the fact that it crashed. This chance is very high.

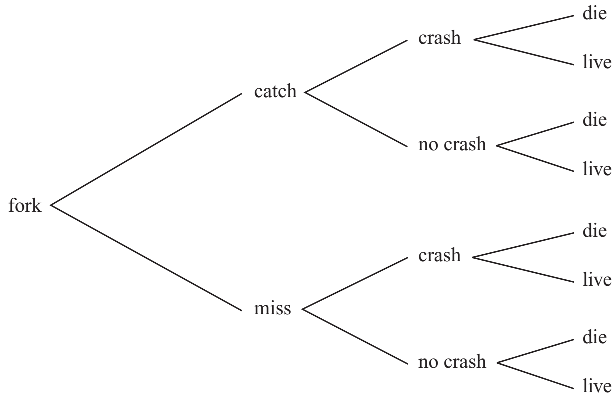

Figure 2.1

This is what we aim at. Of course, we often don't know enough to get it right. Interesting questions arise about how best to estimate such a thing, in states of imperfect information. I won't go into that here.

What happens to the pleasing view of the relation between forward-looking 'will'-conditionals and retrospective 'would'-conditionals, on this amended view? With hindsight, I think that if I had caught the plane, I would have been killed, that if I had bet on heads, I would have won. But there was no reason to think beforehand that if I catch the plane, I will be killed, or if I bet on heads, I will win.

Consider this, however: about the plane crash, a friend has a powerful hunch, or has some erroneous reasons, for thinking this plane will crash. 'Don't take it!' he says. 'If you catch that plane, you'll be killed.' I shrug this off as irrational advice, which it is. I miss the plane. It crashes. 'My goodness, he was right!', I say, on hearing the news. 'If I had caught the plane, I would have been killed!' Similarly, if someone tells me that if I choose ticket number 65 87 92 … I will win, or that if I bet on heads, I will win, or that if I buy these shares, I will become rich, and so on. Even if not rationally grounded, these unfulfilled conditionals are vindicated. The case for the temporal relation between 'wills' and 'woulds' remains: one is right iff the other is. The hindsightful counterfactual vindicates the earlier 'will', even if the 'will' was not justified at the time.

We are familiar with the thought that rationally held beliefs may turn out false and, conversely, something which there is no reason to believe may turn out true. 'Right belief' admits of two readings: rational belief and true belief. If that were my story, there would be no novelty or mystery. But that is not my story. Counterfactuals, like other conditionals, are believed to the extent that a certain conditional probability is judged to be high, and that is not the probability of the truth of a proposition. The right value to assign to them is given by a certain conditional probability, not a truth value. It may be 1 or 0, but it need not be. The fraudulent fortune-teller, gazing into her crystal ball, says 'It's not altogether clear, but I'm pretty sure that if you fly this week, you will be killed.' I miss my plane. It crashes. About 90% of those on board are killed. 'My goodness, she was right!' I say. 'It was very likely that I would have been killed, had I caught that plane.' Lucky guesses are sometimes right, and this was one. The value to be assigned to the hindsightful counterfactual trumps the most rational value to be assigned to the forward-looking indicative. The chance that *C* given *A*, beforehand, provides the best available opinion on whether *C* if *A*, but it can be overturned by subsequent events, not predictable in advance.

What are counterfactuals for?

The question is pressing. Why do we evaluate counterfactuals the way we do? What would go wrong for us if we chose to evaluate them in some other way, for example according to the 'standard picture'? The question deserves more attention than it has had in the vast literature on counterfactuals. I don't pretend to an exhaustive answer, but highlight some important aspects of their use.[3]

We use counterfactuals in empirical inferences to conclusions about what is actually the case. We need to try to get them right, in order to avoid, as much as possible, arriving at wrong conclusions about what is the case. I shall concentrate on two such forms of inference. There may be more, but these, I think, are central. Some examples:

(1a) You are driving, of an evening, in the dark, close to the house of some friends, and have considered paying a visit. You turn the corner. 'They're not at home', you say, 'for the lights are off. And if they had been at home the lights would have been on.'

(1b) 'It's not a problem with the liver', says the doctor, 'for the blood test was normal. And if it had been a liver problem, it would have been [such-and-such].' Call these types of inference 'counterfactual modus tollens'.

(2a) A patient is brought to hospital in a coma. 'I think he must have taken arsenic', says the doctor, after examination, 'for he has [such-and-such] symptoms. And these are just the symptoms he would have if he had taken arsenic.' (Note that in calling the conditional in this inference 'counterfactual' we are using the label as a proper name for a form of conditional. There is nothing literally counterfactual about it. The example comes from Anderson, 1951.)

(2b) The prison warden on his rounds says 'I think a prisoner escaped from that window, for the flowers below are all squashed. And they would have been squashed if he had jumped from there.' Call this style of inference 'inference to a good explanation'.

So we have two forms:

1 H. Because, E; and if it had not been the case that H, it would not have been the case that E;
2 H. Because, E; and if it had been the case that H, it would have been the case that E.

Neither of these forms of inference is valid. They are defeasible forms of empirical reasoning. This is obvious in the case of (2). (2a) could be defeated by pointing out that although these are indeed the symptoms he would have if he had taken arsenic, they are also the symptoms he would have if he had not taken arsenic but was, say, epileptic. (2b) could be defeated by pointing out that the flowers would also have been damaged if a prisoner had not escaped but there had been a game of football, or a dog fight.

The same is true of (1). (1) is closer to being valid in the following sense. If each premiss is certain, the conclusion is certain. But contingent conditional premisses of this kind are rarely certain, and we need to use them when they are less than certain. And an argument deserving the appellation ‘valid’ is such that if both premisses are close to certain, so is the conclusion (not quite so close, perhaps, but still close). That is a property demonstrably had by all paradigmatically valid arguments. (1) does not have this property. It can be defeated thus: ‘I agree that it was indeed very likely that we would find the lights on, if they were at home; but it was also very likely that we would find the lights on, if they were not at home; for they have the deeply engrained practice of leaving the lights on when they go out at night. So there must be some other explanation for the lights being out. Perhaps there’s a power cut; or they have gone to bed early.’

When we see what defeats them, we see that the two forms of inference are not really distinct; in giving one, there is a tacit appeal to the other. To say they are distinct would be like saying that there are two forms of explanation of actions, one in terms of belief, one in terms of desire. ‘He took his umbrella because he thought it was going to rain.’ ‘She went to London because she wanted to see Mike.’ The former could be defeated by pointing out that he loves getting soaked by rain; the latter could be defeated by pointing out that she knew very well that Mike was in America.

So we have:

1 H. Because E. And (probably) $\neg H \Rightarrow \neg E$.
Defeated if (probably) $H \Rightarrow \neg E$.
Undefeated if (probably) $H \Rightarrow E$.

2 H. Because E. And (probably) $H \Rightarrow E$.
Defeated if (probably) $\neg H \Rightarrow E$.
Undefeated if (probably) $\neg H \Rightarrow \neg E$.

That is, for a good argument from E to H, we want it to be probable that if H had not been the case, E would not have been the case; *and* we want it to be probable that if H had been the case, E would have been the case.

These facts are captured by a time-honoured principle of probabilistic reasoning, a form of Bayes’s Theorem:

$$\frac{p_O(H)}{p_O(\neg H)} \times \frac{p_O(E \text{ if } H)}{p_O(E \text{ if } \neg H)} = \frac{p_O(H \text{ if } E)}{p_O(\neg H \text{ if } E)} = \frac{p_N(H)}{p_N(\neg H)}$$

The left-hand equation is a theorem of probability theory, applied to a single probability function, p_O.[4] ('O' and 'N' stand for 'old' and 'new' respectively, and represent probabilities prior to learning E, and posterior to learning E.) The right-hand equation represents the recommendation that on learning E (and nothing else of relevance) your new probability for H should be your old probability for H if E (which, ceteris paribus, is reasonable – some think of this as 'probabilistic modus ponens'). Eliminating the middle term, the equation shows how, on learning that E, your new relative values for H depend on your old together with these conditional factors. In our first example, E is 'the lights are off' and H is 'they're at home'. The inference to ¬H is a good one if it's unlikely that the lights would have been off, if they were at home; unless it is also unlikely that the lights would have been off, if they were not at home.

The equation makes clear that another way the inferences may be defeated is by pointing out that the hypothesis in question *was* very unlikely, before the new evidence: 'but they're always at home at this time'; or 'but they promised they would be in: there must be some other explanation for the lights being off'.

Principles like the above are sometimes called principles of 'updating': they tell you how to 'update' your degree of belief in H, from old to new, in the light of new information E. They can, and do sometimes, have this use. But far more prevalent are instances of their use which involve 'backdating' ('downdating' doesn't sound quite right). To use them in the updating way, you already have to have foreseen the possibility of the information you receive by perception or testimony; and already, before acquiring it, have a judgement about how likely it is that you will acquire it, under various hypotheses. But we continually see, hear, read in the newspaper, and so on, things which we did not anticipate the possibility of coming across. If an observation strikes you as in need of explanation, or as the possible basis of an inference relevant to your concerns, you start there, and ask yourself: how likely *was* it that I *would* get this information, if H? And, if ¬H? Your present 'would haves', as I said before, record your present opinion about the acceptability of an earlier 'will' (an earlier 'will', incidentally, which may concern a time before you were born).[5]

Thus, we need, for the empirical inferences we make, not only judgements to the effect that such-and-such *is* (more, or less) likely; but judgements that such-and-such *was* (more, or less) likely, or likely given something else – that it was more or less likely that it *would* come about, given various hypotheses. We do not fully characterize a person's epistemic state by their present degrees of belief. One could not do much by way of empirical inference without judgements that what I am now certain does obtain, on the basis of my senses, *was* unlikely to obtain, on certain hypotheses, and likely on others. And of course, we should do our best to get such judgements right.

But if this is what we do, in explaining and drawing inferences from what we see and hear, of course we will use hindsight. My final example to illustrate this is similar to the plane crash (and is inspired by Alvin Goldman, 1967). A long time ago, a volcano erupted. It was a slow eruption, the lava creeping onwards slowly.

At that time, it was very likely that the lava would eventually submerge valley *A*, but valley *B* would not be affected – given the lie of the land. However, in the unlikely event of an earthquake of a particular kind at an appropriate time, the path of the lava would very probably be switched away from valley *A*, towards valley *B*. As a matter of fact, this is what happens.

Along comes our geologist, centuries later, making his inference about the eruption. He has already found out about the earthquake. 'That volcano must have erupted', he concludes, 'for there is lava in valley *B* and not in valley *A*; and, given what I know about the earthquake, that is just what one would expect to find if that volcano had erupted.' Also, someone who, before the eruption, said 'If that volcano erupts, valley *B* will be submerged', was unjustified, but, in the event, right.

The point of this example is that *our inferential practices would not be well served* by rejecting counterfactuals which can only be got right with hindsight. Suppose there was a second volcano whose potential eruption, at the time in question, presented much more danger to valley *B*, but in the unlikely event of the earthquake, its lava would probably be diverted elsewhere. Only with hindsight (knowing of the earthquake) is one justified in thinking that if the second volcano had erupted, valley *B* would not have been submerged; and if the first had erupted, it would have been submerged. And it is our hindsightful judgements that stand most chance of leading us to true beliefs. This explains why our practice in evaluating these problematic counterfactuals is as it is.

Given their crucial use in empirical reasoning, then, we see why the 'standard picture' was wrong. We need to take into account actual facts concerning times later than the antecedent-time. We see also, I think, why counterfactuals are best assessed probabilistically. A true/false cut-off point would not serve us well. What matters, for the empirical inferences we make, is how likely it was that E would have happened if H, compared with how likely it was that E would have happened if ¬H.

We have seen a way in which our counterfactual judgements explain and justify our other beliefs. Of course they play other roles. As is implicit in several of my earlier examples, they also explain and justify our reactions of being glad or sorry, relieved or regretful, that such-and-such has happened. 'I'm sorry that Fred didn't come this evening; for if he had come we would have had a fourth for bridge.' This is the retrospective version of 'I want Fred to come this evening; for if he comes, we'll have a fourth for bridge'. (These cases are discussed in Adams 1998.) These positive and negative reactions to what has happened are an important part of our lives, and are assessable as reasonable or not. It is hard to believe that many of our desires, beyond the most basic hard-wired ones, would survive if we were always indifferent to what has happened. In the problem cases where the rational attitude to the forward-looking 'will' differs from that of the retrospective 'would', our reactions switch. I want to catch that plane. If I don't, I'll be late for the meeting. I am dismayed by missing it. On learning that the plane has crashed, my dismay switches to relief: if I had caught the plane, I wouldn't have made the meeting, or any other meetings.

I am spotted in Paris arriving, very late, for the meeting. They had just heard the news. Surprise! ‘She must have missed that plane’, they say. ‘If she had caught that plane she would be dead.’ We’ll leave open whether this is accompanied by relief or disappointment.[6]

Notes

1 At least if the story is told in an appropriate way. As with the plane crash, the betting story is sensitive to whether my saying ‘Heads’ might have influenced the manner in which the coin was tossed. To avoid this possibility, let the tossing happen in one room, and I write ‘Heads’, ‘Tails’ or ‘No bet’ on a piece of paper in another room.

2 Here I borrow from David Johnson (1991), one of the few discussions of this problem.

3 These thoughts owe a great deal to Ernest Adams (1975) Chapter 4, and (1993). No one else, to my knowledge, has investigated this aspect of the use of counterfactuals.

4 Note: equations need numbers. That is an idealization, in examples like those under discussion. But it is a useful idealization. I am not interested in exact values, but only in orders of magnitude: ‘close to 1’, ‘close to 0’, ‘around 50–50’ and the like.

Note also: I have written ‘if’ where probability theorists have ‘given’.

The proof of the left-hand equation is as follows. p(H & E) = p(H) × p(E given H) [Basic Principle]. As p(H & E) = p(E & H), p(H & E) also equals p(E) (H given E). Equating the two longer expressions, we have $p(\text{H given E}) = \frac{p(\text{H}) \times p(\text{E given H})}{p(\text{E})}$.

Call this equation a = b. Derive a similar equation for p(¬H given E). Call it c = d. Then the equation in the text is $\frac{a}{c} = \frac{b}{d}$. Note that p(E) cancels out.

5 The standard literature on probabilistic reasoning ignores the point I am stressing here. It invites the picture of reasoners as ‘probabilistic machines’, which attach values to all the propositions in their repertoire at all times, the values being ‘updated’ as new data are fed in. This is a wildly unrealistic picture, as well as a depressing one.

6 Versions of this paper have been presented at various seminars and meetings over the last few years, including the very enjoyable workshop on chance and causation, organized by Phil Dowe, which led to this volume. I am grateful to many people for criticism, and owe special thanks to Hartry Field for his comments at a seminar at NYU, to Jonathan Bennett and Scott Sturgeon for discussion of these issues, and to Ernest Adams for much inspiration.

Chapter 3

Chance-lowering causes

Phil Dowe

In this paper I reconsider a standard counterexample to the chance-raising theory of singular causation. Extant versions of this theory are so different that it is difficult to formulate the core thesis that they all share, despite the guiding idea that causes raise the chance of their effects. At one extreme, 'Humean' theories – which can be traced to Reichenbach – say that a particular event of type C is the cause of a particular event of type E only if P(E|C & K) > P(E|~C & K) where K is a set of background conditions and where the probabilities are interpreted as relative frequencies. At the other extreme, explicitly non-Humean theories take chance to be a physical, particular, local feature of the world. Mellor, for example, holds that a particular fact C causes particular fact E only if $ch_C(E) > ch_{\sim C}(E)$ in circumstances S, which is to be read as 'the chance that C gives E is greater than the chance of E without C in the same circumstances' and where the chance that C gives E is a local fact about C in S, given by the chance of E in the closest possible worlds in which C and S are true (Mellor 1995).

The obvious counterexample is the type of case where a particular cause lowers the chance of its effect in the circumstances. To my knowledge the first example was posed by Deborah Rosen: a golf player slices her shot, thereby lowering the chance of a hole-in-one, but the ball hits a tree branch and rebounds on to the green, and into the hole for a hole-in-one.

This type of counterexample assumes that there is a method, independent of ascertaining probability relations, for deciding whether one event causes another. Most commonly the method (implicitly) used is intuition: 'intuitively we think that the sliced shot is the cause of the hole-in-one in this case'. So one strategy for defending the chance-raising theory of causation, the 'despite defence', denies the validity of whatever intuitions there might be, and pronounces that the event is not the cause of the 'effect', which instead occurs 'despite' the alleged cause. For those who take the task of philosophy to be conceptual analysis – that is making explicit the concepts inherent in our talk and thought – the despite defence is particularly hazardous. On the other hand, those who take the task of philosophy to be something other than conceptual analysis can define causes as chance-raisers and ignore common-sense intuitions.

I will not be discussing the despite defence (I have dealt with it in Dowe 2000b:

chs 2, 7). I wish to deal instead with an alternative response – where one instead denies that the putative counterexample really is chance-lowering. I will consider three versions of this response, first categorized by Salmon (1998: ch. 14) and show that each fails.[1] I then outline my own path-specific solution, which provides both a diagnosis and a solution of the chance-lowering counterexamples, but which does not save the chance-raising theory of causation. In the final section I offer an independent argument for the solution, based on the intrinsicality of causation.

Counterexample #1

Suppose that a gunman enters the room with a gun, and is about to shoot your friend who stands on the other side of the room. To save your friend, you pull out your own gun, and shoot the gunman just as he is about to shoot your friend. Unfortunately, your bullet passes through the heart of the gunman, through his body, missing all bones, and continues on across the room, into your friend's skull, and kills her. Your firing your gun at the gunman (C) lowered the chance of your friend dying (E). It was very likely that he would have killed her since his gun was loaded, working, he fully intended to do it, he was a good shot, and she stood still not 5 metres away. Your firing your gun was unlikely to kill her since the gunman stood between you and her and the bullet was most unlikely to pass through the man's body. Moreover, since the gunman hasn't seen your gun, you are a good shot, and the gunman stands about 5 metres away, there is a good chance that you can kill him before he shoots. So your firing your gun lowered the chance of your friend dying even though it caused her death as it happened.

It follows that $P(E|C \& K) < P(E|{\sim}C \& K)$ where K includes obvious relevant background factors such those given in the previous paragraph, and for the frequentist this is made true by the fact that in similar situations the victim's life is more often saved than ended this way. It also follows that $ch_C(E) < ch_{\sim C}(E)$, because the chance that my shooting gives E is greater than the chance the same circumstances without my shooting gives E.

There are various strategies to save the chance-raising theory of causation, short of denying the intuition that my shooting caused her death. In this paper I will consider three, which, following Salmon, I will label as follows: (1) 'Fine-grain the cause', where one more closely specifies the cause event, for example, as my shooting the gun in precisely the direction that I did; (2) 'Fine-grain the effect', where one more closely specifies the effect, for example, as her being killed by a bullet of the type that belongs to my gun (and not that of the gunman); and (3) 'Interpolating causal links' where one identifies an intermediate event D between the cause and the effect such that there is a chain of chance-raising C–D–E, for example, accounting for the bullet emerging from the gunman. In what follows I will consider each of these strategies in turn, providing versions of the counterexample which are not amenable to these strategies.

Fine-grain the cause

If we more adequately specify the cause, according to this strategy, we will find that the cause does raise the chance of the effect. Take C' to be my shooting in exactly the direction that I did. Then $P(E|C' \& K) > P(E|{\sim}C' \& K)$ simply because my shooting in that exact direction, unknown to me, did make it very likely that the bullet would pass through the gunman. For the same reasons, Mellor's approach gives $ch_{C'}(E) > ch_{\sim C'}(E)$ (although the strategy is not available to a factualist, because that I fired the gun and that I fired the gun in the direction I did are distinct facts).

We need to note that in general it is necessary to fine-grain the circumstances to the same extent as the cause. It will not do to specify the cause as C' if the background K is specified only roughly. For example if we think of K including the fact that the gunman stood between me and the victim, this in itself may not be a fine enough description to give the desired chance relations. We need to specify exactly where and how he stood in order to make it sufficiently probable that the bullet would pass through. This fine-graining of the background is already built into Mellor's account, where the circumstances S include all the local facts that there are.

We should next ask, to what degree should we fine-grain the cause? What facts about my shooting should be included in the cause? We have already seen that the chance relations can be reversed by more closely specifying the cause, and indeed in principle they can be reversed back again by yet more closer specification. And the question is not only 'how far should we fine-grain?' but worse, 'why isn't our answer to that question arbitrary and question-begging?' It would be arbitrary if we have no reason to prefer the adopted degree of fine-graining to any other, and question-begging if we choose a level of fine-graining just because it gives the desired result.

One obviously non-arbitrary answer would be always to fine-grain completely – in other words, specify everything that is true about the cause, say C* (with an appropriate restriction to local factors, or the like, so as to exclude properties such as 'eventually has effect E').

However, this answer faces a dilemma. Either our situation is deterministic or it isn't. By deterministic I mean that the state of the world at the time of the cause, together with the laws of nature, fixes the state of the world at the time of the effect, and, conversely, the state of the world at the time of the effect, together with the laws of nature, fixes the state of the world at the time of the cause. This means that $P(E|C^* \& K^*) = 1$ where K* includes everything (local) about the situation in which C occurs at the time of C, and $ch_{C^*}(E) = 1$, if C is what Mellor calls a 'total cause' (if it is not, then the conjunction of C* with whatever else makes up the total cause gives E a chance of 1).

Suppose on one hand the situation is deterministic. Then $P(E|C^* \& K^*) = 1$ and $ch_{C^*}(E) = 1$. But we also need to know $P(E|{\sim}C^* \& K^*)$ and $ch_{\sim C^*}(E)$. Take the frequency version. Here the problem is that it may well be that the conjunction of

~C* & K* is physically impossible, in other words that P(~C* & K*) = 0. The reason is that, in a deterministic world, factors which are part of C* may have as a cause something which is also a cause of a factor which is part of K*. This may be true of all parts of C*, in which case P(~C* & K*) = 0. But if P(~C* & K*) = 0 then by the definition of conditional probability P(E|~C* & K*) is not well defined.

Consider instead the counterfactual chance $ch_{\sim C*}(E)$. Given that the situation is deterministic, then this will be either 1 or 0, depending on which of the alternatives to C is found in the closest world. If I don't shoot is the gunman sure to kill her, or is he sure to miss? If he is sure to miss then it seems that we have saved chance-raising, because 1 > 0. But perhaps he is sure to kill her; that is, $ch_{\sim C*}(E) = 1$ (deterministic preemption). Then we have not saved chance-raising, because 1 = 1. And since it is contingent whether this chance is 1 or 0, the strategy fails.

Suppose on the other hand the situation is relevantly indeterministic, such that P(E|C* & K*) < 1 and $ch_{C*}(E) < 1$. Then it is possible that we have chance-lowering, without there being any relevant further factors that will enable us to fine-grain the cause. This is not the situation with the present example, but below I will provide a case where it is.

So it seems whether the situation is deterministic or indeterministic the strategy fails. We cannot even restrict its applicability to one or other kind of situation.

Fine-grain the effect

This strategy develops the idea that effects brought about by alternative causes are really different effects, and draws on the possibility that there will always be some difference. In our example, suppose my gun and that of the gunman are quite different – for simplicity suppose than my gun uses silver bullets and his uses lead. Then we introduce E' (killed by a silver bullet) and E" (killed by a lead bullet). Then since in fact she was killed by a silver bullet, we replace E with E', and note that P(E' |C & K) > P(E' |~C & K) in keeping with our intuition that I caused her death. Similarly, $ch_C(E') > ch_{\sim C}(E')$. Sure I prevented E" by my action, since P(E"|C & K) < P(E"|~C & K), but that is a different death to the one I caused. (A version of this approach uses the times of effects to disambiguate. Then, providing alternative possible causes would occur at different times, the problem is again avoided (Paul 1998b).)

This strategy assumes that alternative possible causes would leave different traces, a thesis due perhaps to Leibniz. If true this would allow us to in principle distinguish alternative effects. But if this assumption is true, then it is so only contingently. For a start it doesn't seem to involve any contradiction to suppose there are alternative possible effects that are totally indistinguishable. Indeed, I will later give an actual example of indistinguishable alternative effects.

Interpolating causal links

In the third strategy one identifies one or more events between the cause and the effect such that there is a chain of chance-raising even though there is no chance-raising between the end points. One can then define causation as the ancestral of chance-raising (in the terminology of Lewis 1986: ch. 21): two events (C, E) are cause and effect only if either there is chance-raising between them; or there is a third event D such that C raises the chance of D and D raises the chance of E; or there are third and fourth events (D, F) such that C raises the chance of D, D raises the chance of F and F raises the chance of E; or … and so on. Alternatively one can define direct causation as chance-raising (an intransitive relation) leaving the possibility of a relation of indirect causation (a transitive relation) as the ancestral of direct causation. In our example, call B the bullet emerging from the gunman's heart. Then my shooting certainly raises the chance of B, and B raises the chance of E (her death), since we need to include in this latter relation (as part of background conditions) the fact that the bullet has entered the gunman's heart, killing him before he can shoot his gun.

So $P(B|C \,\&\, K) > P(B|{\sim}C \,\&\, K)$; and $P(E|B \,\&\, K') > P(E|{\sim}B \,\&\, K')$ and similarly $ch_C(B) > ch_{\sim C}(B)$ and $ch_B(E) > ch_{\sim B}(E)$. For the last relation to hold it is required that the closest worlds without B include the gunman's death – a semantics such as that of Lewis would do the trick where the closest worlds have perfect match with the actual world up to the time of B, then by a small miracle B fails to occur.

The problem with such a strategy is simply that there is no guarantee that there will be such an event between the cause and effect. This becomes particularly pressing if time is discrete, so that one could have the cause occurring at one instant and the effect occurring at the immediately following instant, and where we have chance-lowering. I will give an example of such a case on p. 33. I am not sure that we have discrete time in the actual world – I myself find arguments from quantum mechanics to be inconclusive – but it seems to be a distinct possibility.

Counterexample #2: the decay case

We have considered three strategies for dealing with the chance-lowering counterexample. The first, 'fine-grain the cause', was successful in dealing with Counterexample #1, provided the situation is not deterministic and such that had I not fired my gun the gunman would have been certain to kill the victim. But I promised an indeterministic example where there are no further features to allow the fine-graining strategy to operate. The second, 'fine-grain the effect', was also successful in dealing with Counterexample #1, but I promised an example where alternative possible effects are indistinguishable, which would thwart this second strategy. The third, 'interpolating causal links', was also successful in dealing with Counterexample #1, but I promised a counterexample where time is discrete and the chance-lowering cause and effect occur at immediately adjacent instances, thereby thwarting this third strategy. It's now time to come good on these promises.

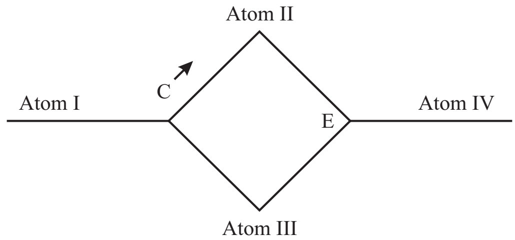

Figure 3.1

This counterexample (originally due to Salmon, and presented in Dowe 2000b: 33–4) involves a case of nuclear decay. Atom I can decay by two possible routes to atom IV. One, decay via atom III, is unlikely to lead to IV because atom III is more likely to decay to some other product. The other, via atom II, is likely to lead to IV because atom II is unstable and can decay only to IV.

Suppose we have the following transition probabilities:

P(I–II) = 0.5
P(I–III) = 0.5
P(II–IV) = 0.1
P(III–IV) = 1

In our particular case we get the decay sequence I–II–IV. We want to say that the decay of atom I to II is a cause of the production of atom IV. Writing C for the decay of atom I to II and E for the production of atom IV, we have P(E|C & K) = 0.1 < P(E|~C & K) = 1, taking ~C to be the class of events such that atom I decays to III. It also follows for similar reasons that $ch_C(E) = 0.1 < ch_{\sim C}(E) = 1$, again, the closest ~C-worlds are worlds in which atom I decays to III.

If we take the time of C to be essential to C, then at least on Lewis's similarity relation it is not clear to me what $ch_{\sim C}(E)$ is, since I do not know whether the closest ~C-worlds would contain, at the time of C, atom I undecayed and $ch_C(E) = 0.55$ or atom I decaying to II and $ch_{\sim C}(E) = 1$. But either way we have chance-lowering.

The strategy of fine-graining the cause cannot work here, because, besides the time of the cause, there is no further fact about the decay which is going to change the chances. There is no hidden feature about atom I, atom II or the decay itself which changes the chance that atom II will decay to atom IV. That chance is irreducible (for discussion of this see Dowe 2000b: 24–5).

The second strategy, fine-graining the effect, works because E is defined as the production of IV, which covers two alternatives: the decay of II–IV and the decay of III–IV. Let E' be the decay of II–IV and E" the decay of III–IV. Then P(E'|C & K) = 0.1 > P(E'|~C & K) = 0.

However, this strategy will not work if we modify the example slightly. Take F to be the existence of the IV atom after production. Intuitively in our case C is the cause of F. Again, $P(F|C \& K) = 0.1 < P(F|{\sim}C \& K) = 1$, and $ch_C(F) = 0.1 < ch_{\sim C}(F) = 1$. But now there is no feature about F which will distinguish the two ways it may have come about. Suppose the upward paths involve an alpha decay and the downward paths a beta decay. The end product of the I–II–IV decay (a IV atom, one alpha particle and one beta particle) is identical to the product of the I–III–IV decay. Perhaps it is possible that there are different energies involved, but it is also possible that all the energies are identical. Providing that we may only include local, intrinsic facts about F, there is nothing to distinguish the alternative possible effects, so the strategy fails.

The third strategy, interpolating causal links, fails if we modify the counterexample in another direction. Suppose we have discrete time, and at t_1 atom I exists undecayed, at t_2 C occurs, and at t_3 E occurs, where t_1, t_2 and t_3 are successive instances. Elsewhere I called this ‘cascading decay in discrete time’ (Dowe 2000b). Then we have a chance-lowering cause as above, but there is no instant between the time of the cause and the effect at which an event can occur which would allow this strategy to operate.

So each of the three strategies fails, for different reasons. This does not mean that a combination would fail, for example, such that we can use the ‘interpolating causal links’ strategy to deal with the case for which the ‘fine-grain the effect’ strategy fails, and the ‘fine-grain the effect’ strategy to deal with the case for which the ‘interpolating causal links’ strategy fails.[2] But we have shown that each strategy by itself fails, which is enough to motivate the solution provided in the next section.

Counterexample #2 also shows, if it is successful in thwarting the above three strategies, that these strategies are not really a satisfactory answer for those counterexamples – for example, Counterexample #1 – for which they do provide the right answer. This point is demonstrated by the following diagnosis of chance-lowering causes, which also explains why the three strategies work to the extent they do.

The path-specific solution

I claim the reason we have chance-lowering in our counterexamples is not because we have failed to adequately specify the relata-events, nor because there are intermediate events that we have failed to account for. In my view the reason is that in such cases there are two possible causal paths between the cause C and the effect E, and via one path C tends to cause E and via the other path C tends to prevent E.[3] Further, the causing path is the ‘weaker’ or less reliable of the two, yet is successful, while the preventing path is not successful in that it fails to prevent E, even though it is the ‘stronger’ of the two paths. (More will be said on p. 36 about what is meant by ‘weaker’ and ‘stronger’.[4]) This explains why we have a chance-lowering cause. The cause, as well as initiating a causal process that leads to the effect, also preempts a stronger process which, if successful, would have prevented the effect (see Dowe 2000b for more details). My hypothesis is that all counterexamples involving chance-lowering causes have this structure.

We can now show in detail how chance-raising and causation stay together despite our counterexamples. (The following nine points summarize the path-specific solution given in Dowe 2000b: 165–7 and Dowe 1999: S498–500. For criticism see Beebee, this volume.)

1 A cause and its effect can be linked along more than one path. For example, I fire my gun at you, the bullet passes through a rope, which causes a large rock to fall on your head just as the bullet enters your skull, over-determining your death. In this case both paths are causal paths.
2 Two paths between a cause and its effect may be ‘opposed’. By this I mean that one of the paths is a causing path, and the other a preventing path. These paths are delineated by actual causal processes and by preventing paths (which I take to be a matter of the possibility of causal processes). That they cannot be delineated spatiotemporally is illustrated in Counter-example #1, if we suppose that the bullet from the gunman would have occupied the same spatiotemporal locations as did my bullet, together with the fact that the first section of the two paths is identical.
3 When a cause and its effect are linked by two opposed paths, only one can be successful, since you cannot get both E and not E. In our counterexamples there are two paths, opposed, and the causing path rather than the preventing path is the successful one. My firing my gun is linked to the effect via a successful causing path which follows the path of my bullet through the gunman and on into my friend’s head; and via an unsuccessful preventing path, where my bullet kills the gunman, preventing him shooting my friend.
4 Chance has components, in some respects like the components of force, which combine in some way to give the total chance. For example, in Counterexample #1 the chance my firing my gun gives my friend’s death has two components which combine to give the total chance $ch_C(E)$, one in virtue of the fact that my bullet might go on to kill her, the other in virtue of the fact that my bullet might prevent the gunman from shooting.
5 In cases of chance-lowering causes, via the successful path ‘in itself’ the cause C raises the chance of E, and via the other path ‘in itself’ the cause C lowers the chance of E. In virtue of the successful path from C to E ‘in itself’ – following my bullet through the gunman’s body and on into the head of the victim, killing her – my shooting raises the ch(E). But in virtue of the other path from C to E ‘in itself’ – following my bullet to the gunman’s heart, killing him, and preventing him firing his gun – my shooting lowers the ch(E).
6 The path-specific chance relation between C and E is given by the chance relation between C and E in the closest worlds in which that path is the only path between C and E. Take worlds in which there was no way the victim could have died except by my bullet. The $ch_C(E)$ in that world is the component $ch_C(E)$ in virtue of the successful path, and in that world $ch_C(E) > ch_{\sim C}(E)$.

7 The actual chance relation $ch_C(E)$ is the 'combination' (in some sense) of the path-specific component chances. We need not specify exactly what mathematical operation captures this 'combination'.
8 We can now explain what we mean by 'stronger' and 'weaker' paths. If the components add together to give a total chance-raising relation between C and E, then the causing path is stronger, and if the components add together to give a total chance-lowering relation between C and E, then the preventing path is stronger.
9 We can say causes raise the chance of their effects if we take the relevant chance to be the path-specific chance relation along the successful path – the actual causal process linking the cause and effect.

To illustrate how the path-specific account works, consider our two counter-examples. In the first, we may say that my shooting raises the chance of my friend's death because in the closest world without the possibility of the gunman shooting my friend, my shooting raises the chance of her death. In the second, in the closest world where there is only one decay path between atom I and atom IV (via atom II), the decay of atom I raises the chance of the production of atom IV.

The path-specific account shows how causation and chance-raising 'stay together', but it will not save the chance-raising analysis of causation, because it assumes we know which path is actually causal. In *Physical Causation* (Dowe 2000b) I provide an independent account of causation, which does not utilize the concept of chance-raising.

Now we can see why the three strategies considered above work to the extent that they do. The 'fine-grain the cause' strategy selects factors which make the successful path the stronger. The 'fine-grain the effect' identifies factors which indicate which path was successful and takes the chance-raising relation specific to that path. The 'interpolating causal links' approach identifies an event which is part of the successful causing process, thus capturing chance-raising relations specific to that path.In the final section I offer a further argument for why the path-specific account should be adopted by chance-raisers (that is, people who think that causes raise the chance of their effects).

An argument from intrinsicality

An increasing number of philosophers are appealing to the (alleged) intrinsicality of causation, in other words the intuition that whether a process is causal is an intrinsic matter. Lewis writes:

> Intuitively, whether [a] process going on in a region is casual depends only on the intrinsic character of the process itself, and on the relevant laws. The surroundings, and even other events in the region, are irrelevant.
>
> (Lewis 1986: 205)

Peter Menzies, in his paper 'Intrinsic versus Extrinsic Conceptions of Causation' writes:

> The causal relation does not depend on any other events occuring in the neighborhood: the causal relation is intrinsic, in some sense, to the relata and the process connecting them.
>
> (Menzies 1999: 314)

Elsewhere Menzies takes intrinsicality to be a fundamental platitude in the folk concept of causation (Menzies 1996). And David Armstrong writes: 'The causal structure of a process is determined solely by the intrinsic character of that process' (Armstrong 1999: 184).

The spatiotemporal relations between the two paths are contingent. In Counterexample #1 the two paths coincide substantially. But they need not. Suppose I push my friend out of the way of a bus, on to the footpath, but unfortunately she is hit by a falling stone and dies. If the bus was more likely to kill her than the stone, then we have chance-lowering cause, with the same structure as our examples. But now the two paths are spatially separated. If causation is intrinsic, then it does not depend on what happens or doesn't happen elsewhere, such as on the roadway.

Further, if causation is intrinsic, then it not only does not depend on what actually happens elsewhere, it also cannot depend on what might have but didn't happen elsewhere. But the standard chance-raising relation in our counterexamples depends on what might have but didn't happen elsewhere. It is affected by the mere possibility of other processes occurring between the cause and the effect. The intrinsic intuition tells us that these extrinsic features should not determine the causal character of the actual process. (For a counterargument see Beebee, this volume.)

For example, that the gunman failed to fire the gun seems irrelevant to whether my shot caused the death of my friend – that it is causation is true in virtue of the process running from my shot, the bullet, its progress through the gunman's body and on into the victim's body. The preventing of a decay to atom III is irrelevant to whether the decay to atom II caused the production of atom IV.

So if we accept the intrinsicality intuition, and we want to analyse causation in terms of probabilities then it seems we really ought to reject any idea that the path-specific chance attached to the non-actual causal process is relevant to whether C actually caused E.

Finally, the path-specific solution also explains why some of the above strategies work to the extent that they do. The 'fine-grain the effect' and 'interpolating causal links' strategies both in effect select out the actual process and screen out the alternative path. But they don't follow naturally from the intrinsicality intuition in the way that the path-specific solution does.

Acknowledgement

This work was supported by the Australian Research Council.

Notes

1 The second of these counterexamples is already developed in detail in Dowe 2000b; however, this presentation replaces that account in that it corrects a number of errors, clarifies the argument, and directs the argument at conceptual analysis.
2 An example is the new account of David Lewis (Lewis 2000). I have argued elsewhere that this account fails for other reasons (Dowe unpublished).
3 Here I summarize a point made in Dowe 2000b: 164–5.
4 In counterexamples to the claim that prevention is chance-lowering, the successful path is a weaker, preventing path: see Dowe 2000a.

Chapter 4

Chance-changing causal processes

Helen Beebee

1 Causation, chance increase, and causal processes

Consider the following two prima facie plausible claims about causation:

(CP) Causes and effects must always be connected to each other via a causal process.
(IC) Causes must always raise the chances of their effects.

For the purposes of this paper, I shall not attempt anything like a formal definition of a causal process, but will rest content with a common-or-garden, intuitive conception.[1] For example, there is a causal process between my hitting the white ball with the cue, the white hitting the black, and the black landing in the pocket. There is a causal process between my being in the presence of someone with flu, my getting flu, my gradual recovery (involving my antibodies fighting the infection and so on), and my return, a week later, to full health. There is no causal process between my writing these words, the neighbour's dog barking, and the light in the next room being on. There is no causal process between my failure to go to the supermarket and my subsequent failure to cook dinner.[2] I shall also assume that (IC) is to be read counterfactually: *c* increases the chance of *e* if and only if, had *c* not occurred, the chance of *e* (just after the time at which *c* in fact occurred) would have been lower than it actually was.

As stated, each of (CP) and (CI) claims to represent a necessary, though not sufficient, condition for causation. Some stock examples suffice to show why neither the obtaining of a causal process nor increase in chance should (individually) be taken to be a sufficient condition for causation. First, it is generally agreed that there can be a causal process between *c* and *e* without it being true that *c* caused *e*; the example given above of my getting flu and my subsequent healthy state a week later will do. A structurally similar and well-known case, which I shall call the defoliant case, involves a plant being sprayed with defoliant (*c*), recovering, and eventually being in full health again (*e*).[3] There is a genuine causal process between *c* and *e*, but nobody, so far as I am aware, claims that *c* caused *e*. Second, there can be increase in chance without causation. Fred and Ted both want Jack dead. Fred poisons Jack's soup and Ted, unaware of Fred's act, poisons

Jack's coffee. Suppose that each act increases the chance of Jack's death. Jack eats the soup but, feeling rather unwell, leaves the coffee – and dies later of poisoning. Ted's act raised the chance of Jack's death but did not cause it.

A simple diagnosis of the two kinds of case suggests a straightforward way of providing a sufficient condition for causation. In the flu and defoliant cases there is a causal process between *c* and *e*, but *c* lowers, rather than increases, the chance of *e*: getting the flu decreases my chance of being fully healthy a week later, and being sprayed with defoliant decreases the plant's chance of being fully healthy six months later. In the poisoning example, on the other hand, while *c* increases the chance of *e* there is no causal process between the two: no chain of events or process links Ted's act with Jack's death, since Jack did not so much as go near the poisoned coffee. Hence, by taking the causal process condition and the chance-increase condition to be *jointly* sufficient for causation, we rule out all the problem cases in one neat manoeuvre.

So far so good. Now, if only those conditions were also *necessary* for causation – that is, if only (CP) and (IC) were true – we would have ourselves necessary and sufficient conditions for causation.[4]

First, (CP). (CP) sounds plausible enough, but on closer inspection its truth is by no means obvious. For one thing, it rules out at least some cases of causation by absence, and also prevention (understood as the causing of the absence of an event). For another thing, it rules out causation at a temporal distance, since such causation would, by definition, involve the causing of one event by another without the aid of any process linking the two. (Strictly speaking, it also rules out direct, but not at-a-distance, causation. If time is quantized then, between two events that are as close to each other in time as it is possible to get, there cannot be any further events – so, strictly speaking, there can be no causal process between them. We could circumvent this problem by characterizing a causal process as a process that is 'non-gappy' rather than as a sequence of events between which one can always interpolate further events that hook them together. But I shall leave such technical issues aside.) An adequate defence of (CP) is a big (some will doubtless think impossible) job; but for the purposes of this paper, I shall set such worries aside and assume that (CP) is true.[5]

What about (IC)? Well, there are some alleged counterexamples to it. One alleged counterexample is as follows: Sue pulls her golf drive (*c*), thereby lowering her chance of a hole-in-one. However, fortunately – and against the odds – having hit a tree and bounced back on to the fairway, the ball lands in the hole (*e*). *c* lowered the chance of *e*, but one might still be inclined to say that *c* caused *e*.[6] Another alleged counterexample runs as follows. An old lady – let's call her Edna – is crossing the street, just as a bus comes hurtling towards her. I see the bus, and push Edna on to the pavement, out of the path of the bus (*c*). Unfortunately, however, I push her into the path of a falling brick. The brick is not falling from a great height, and Edna is wearing a sturdy hat; nonetheless, improbably, the brick hits her on the head and she dies. It seems right to say that my push caused Edna's death, even though in pushing her I greatly reduced its chances.[7]A standard move at this point is to claim that Edna's actual death-by-brick is a different event to the event – the death-by-bus – that probably would have occurred had I not pushed Edna on to the pavement. This move restores chance increase, since had I not pushed Edna, her chance of dying the particular death she actually died (rather than some other death) would have been zero. In the current case, this move seems entirely plausible. There is no a priori reason to hold that whenever one attempts, but fails, to save a life, the actual death that occurs is the very same death that would have occurred in the absence of the attempt; the identity of the two deaths may hold in some cases but not in others. Compare, for example, the current case with a case where I push Edna but (say because Edna is very big and I am very weak) simply fail to move her out of the bus's path. In the former case, the push initiates a causal process that is entirely unlike the causal process that would have continued in the absence of the push, and results in a death whose manner is entirely unlike the manner of the death that would have (probably) occurred in the absence of the push. In the latter case, the death that actually occurs and the causal process that leads to it are (perhaps not precisely, but more or less) the same in manner as the causal process and death that would (probably) have occurred in the absence of the push. Plausibly it is only in the latter case that the two deaths are identical. But in the latter case we are much less inclined to say that the push was a cause of Edna's death.

Moves like this are not, unfortunately, always possible. Suppose that the actual brick-death and the non-actual bus-death are indeed two different deaths, but that their precise time and manner are not so different as to affect later consequences. Suppose, for example, that whichever death Edna were to die, her funeral would be conducted at the same time and place, and in the same way. Call the funeral event *f*. Then while *c* does not (if the brick-death and the bus-death are sufficiently different in manner to count as different deaths) lower the chance of *e*, it does lower the chance of *f*: had *c* not occurred, the chance of that very funeral's happening just as it did would have been greater, since Edna's chance of dying (as opposed to dying the particular death she actually died) would have been greater had I not pushed her. For the purposes of simplicity I shall assume that the death-by-bus and the death-by-brick are the very same event. However, everything I say about the relation between the push (*c*) and Edna's death (*e*) might just as well be said about the relation between *c* and *f*, so readers who deny that the death-by-bus and the death-by-brick are the same event should substitute *f* for *e* in subsequent discussion of the case.

If common-sense intuitions about such cases are to be taken seriously (though I shall argue later that they need not), the chance-increase condition cannot be taken to be necessary for causation (which is to say, (IC) cannot be true), since in such cases *c* causes but lowers the chance of *e*.

Alleged cases of chance-decreasing causation create a problem not just for the prospects of providing necessary and sufficient conditions for causation in terms of causal processes plus chance increase; they also create a problem for what Phil Dowe calls the 'Chance Changing Thesis' (hereafter (CCT)):

(CCT) *c* promotes or tends to cause *e* when *c* raises the chance of *e*, and *c* causes *e* when *c* successfully promotes *e* (in other words when *c* raises the chance of *e* and *e* occurs).

c hinders (or inhibits) *e* when *c* lowers the chance of *e*, and *c* prevents *e* when *c* successfully hinders *e*, in other words when *c* lowers the chance of *e* and *e* fails to occur.[8]

(CCT) is an appealing thesis, and one that Dowe wants to uphold. However, since (CCT) implies (IC), counterexamples to (IC) are counterexamples to (CCT) too. He therefore adopts a strategy for dealing with the counterexamples that runs as follows. First, he diagnoses the problem cases – cases of chance-lowering causation – as cases where there is a 'mixed path' from *c* to *e*. Roughly speaking, the idea is this. In mixed-path cases, *c* initiates two 'processes': one that could lead to – or tends to cause – *e* and one that could prevent, that is hinders, *e*. Since the hindering process is stronger than the promoting process, *c* lowers the chance of *e*.

Dowe's solution to the problem of chance-lowering causes is to distinguish between (what I'll call) the 'all things considered' chance of *e* (this is the chance of *e* that *c* lowers) and the chance of *e* 'relative to a process'. He argues that if we abstract away from the hindering process and consider the chance of *e* relative just to the promoting process, then chance increase is restored: relative to this process, the chance of *e* in troublesome mixed-path cases really is increased by *c*. Thus (CCT), and therewith (IC), is saved – once we interpret 'chance' as 'chance relative to a process' rather than the usual 'all things considered' kind.

For example, in the bus–brick case, pushing Edna out of the path of the bus and into the path of the brick initiates two different 'processes'. On the one hand, the push initiates a genuine causal process (in the intuitive sense described earlier) which in fact culminates in Edna's demise. On the other, the push initiates another 'process' that might (and indeed is intended to) save Edna's life. (This 'process' involves Edna's not being in front of the bus as it speeds along, not being hit by it, and so on. This is *not* a genuine causal process in the sense described earlier, since it is a sort of chain of non-events rather than a chain of events.) Unfortunately for Edna, this hindering 'process' is unsuccessful – it fails to prevent her death.

The chance-relativizing thought is roughly this: abstract away from the (hindering) bus-avoiding 'process' and concentrate just on the (promoting) brick process. Intuitively, just imagine that the bus was never there. In such an imagined scenario, the push is not a potential death-preventer, since had Edna remained in the middle of the road she would have been perfectly safe. The brick process in the imagined scenario, however, *is* still there. So in that scenario – which is to say, relative to the brick process – the push really does increase the chance of Edna's death, since had I not pushed, Edna would have been very much less likely to die.

Like Dowe, I want to hang on to (IC). However, as I argue in Section 2, his strategy for saving (IC) is unsuccessful. In Section 3, I present a different conception of 'hindrance', according to which hindrance is a causal relation manifested by chance-lowering causal processes. I show how analyses of causation that do not count hindrance as a bona fide causal relation face a major problem, and argue that, with the appropriate notion of hindrance firmly in place, biting the bullet with respect to alleged cases of chance-lowering causation is a plausible strategy. The bullet-biting strategy simply denies that common-sense intuitions concerning the alleged counterexamples deserve to be taken seriously enough to undermine (IC), and I argue that the desire to believe in chance-decreasing causes can be explained away. Finally, in Section 4, I argue that the alleged cost involved in this strategy – the denial of the transitivity of causation – is no cost at all, since there are no good arguments for the claim that causation is transitive. I also show that another objection of Dowe's – that the strategy makes causal facts depend on facts that are extrinsic to the causal process in question – fails to hit home.

2 Dowe's chance-relativizing strategy

In this section I argue that Dowe's strategy for rescuing (IC) does not succeed. I begin by showing that his analysis of a hindering process – the process from which we need to abstract away in order to restore chance increase between *c* and *e* – fails. However, the failure of that analysis does not entail that we should abandon the general chance-relativizing strategy altogether, and I therefore provide a more intuitive, and more successful, conception of the alternative process from which we need to abstract away. I then show that given this conception of the alternative process, the chance-relativizing strategy can be applied to chance-decreasing causal processes that are *not* intuitively cases of causation, for example the defoliant case described on p. 39. In other words, all cases of chance-decreasing causal processes can be characterized as mixed-path cases, and hence all such cases are, according to Dowe's strategy, cases of causation. So Dowe's strategy is *too* successful: it makes not just some but all chance decreasers come out as causes. Hence the general strategy does not succeed because it fails to discriminate between (alleged) chance-lowering causes and chance-lowering non-causes.

At first sight, Dowe's strategy seems intuitively plausible in the bus–brick case. There, two identifiable and reasonably independent causal processes are going on: the process that involves the bus speeding along the street, and the process that involves the brick falling. My push in effect stops Edna interacting with the first causal process, but at the same time forces her to interact with the second. So – at first sight – it seems plausible to say that the push does two things: it promotes Edna's death by involving her in the brick process, and hinders her death by getting her out of the way of the bus process.

It also seems to be a straightforward matter to imagine the situation with one of the processes removed: it's easy to imagine the situation minus the bus, where I push Edna (for reasons unknown, or perhaps for no reason), the brick falls, and Edna dies; and it's also easy to imagine the situation minus the brick – where I push her out of the path of the bus and safely on to the (falling-brick-free) pavement.

Clearly in the former case, where the bus isn't on the scene, the push increases the chance of Edna's death. Hence it seems plausible to say that the push initiates a 'mixed path' to Edna's death, and that we can sensibly relativize the chance of Edna's death to just one path: the brick process.

However, we need to be a bit clearer about the nature of the 'paths'; in particular, we need to be precise about the nature of the bus-avoiding 'process' from which we are supposed to be abstracting away. I shall argue that the details of Dowe's analysis do not yield the intuitive picture presented above, and hence that if we want to try to save some form of the chance-relativizing strategy, we are going to have to hang on to the intuitive picture rather than the details of the analysis.

Picture the scene. Before the push, the bus is hurtling down the street and heading straight for Edna. This is a bona fide causal process. Then we have the push. (The bus continues to hurtle down the street, but these later stages of the process are no longer relevant to Edna's death.) The push *stops* Edna from interacting with the genuine causal process of the bus's travelling down the street. However, according to Dowe, when we abstract away from the bus-avoiding 'process', we are supposed to be abstracting away from the hindering 'process' *initiated by the push*. We are not – or at least not directly – supposed to be abstracting away from (the earlier stages of) the genuine causal process of the bus hurtling down the street, since *that* process is neither a potential preventer (that is, hinderer) of Edna's death, nor a process initiated by the push.

So in what sense does the push initiate a hindering process – a process that could (but in fact does not) lead to Edna's survival? Well, the push prevents a particular course of events – say, bus-being-a-foot-away-from-Edna (event *b*), bus-hitting-Edna (*d*), … , Edna's death (*e*) – from occurring: without the push, that sequence of events would have been very likely to occur. On Dowe's account, the 'process' *actually* initiated by the push is, as it were, the negation of the earlier stages of that merely possible process *b*–*d*–*e*. Call the bus's hurtling down the street prior to the push *a*, and the push *c*. Without *c*, the process that would have been very likely to occur would have been *a*–*b*–*d*–*e*. *With* the push, the bus-avoiding 'process' that *actually* occurs is $a–c–\neg b–\neg d–e$. (Recall that on Dowe's view hindrance is unsuccessful prevention. Had the brick not been there, the hindering 'process' initiated by *c* would have succeeded – it would have prevented Edna's death. In other words, it would have been $c–\neg b–\neg d–\neg e$. But the hindering 'process' was not successful, since Edna in fact dies, hence $c–\neg b–\neg d–e$.) Note that this hindering 'process' is not a *causal* process according to the conception of causal processes presupposed throughout this chapter; nor is it a causal process according to Dowe.

Now, Dowe's strategy for saving (IC) is to relativize the chance of *e* to the brick process (call this process ϱ). To do this, we need to 'go to the closest [ϱ]-only world, that is to say, the closest world where [ϱ] is the only process between *c* and *e*' (Dowe 2000a: 80). In other words, we need to evaluate the chance of *e* at the closest world where the other bus-avoiding 'process' (call it the σ-process) does not occur, but the ϱ-process does.

What is such a world like? Well, according to the intuitive characterization of the situation given earlier, we can think of the closest ϱ-only world as one where the bus is simply not on the scene: there, there is (intuitively) no hindering initiated by the push, because the push does not save Edna from any potentially life-threatening road accident, but there is still the genuine causal process from the push, via the brick, to Edna's death. However, this intuitive picture is not what is entailed by Dowe's own analysis. For in such a world, b and d do not happen. So – like the actual world – that world is a world where c, $\neg b$, $\neg d$ and e all 'occur'. But that 'process' is precisely the process – the σ-process – from which we are supposed to be abstracting away.[9] So the world in which we intuitively *want* to evaluate the chance of e in order to restore chance increase between c and e is *not* a world where the σ-process does not occur. Indeed, the only way of getting to a world where $\neg b$ and $\neg d$ do not occur is to go to a world where b and d *do* occur – that is, a world where Edna *does* get hit by the bus. In such a world, pushing Edna fails to get her out of the path of the bus, and hence, so far as I can tell, makes no difference to her chance of dying.

One reason why Dowe's account fails, then, is that the process we need to abstract away from is not the 'process' – the σ-process – he *says* we need to abstract away from. We really need to go to a world where there is no possibility of Edna getting hit by the bus, for it is only in such a world that the push will fail to be a potential preventer of her death. There are three different kinds of world that satisfy this requirement. The first is a world where early stages of the causal process involving the bus – stages that occur before the push occurs – do not happen, for example a world where there is no bus at all. The second is a world where earlier stages of the bus process *are* present, but there are other features of the situation that make it impossible for the process to run to completion. One such world would be a world where there is a sufficiently sturdy barrier – a reinforced concrete wall, say – between Edna and the bus. The third and final kind of world is a world where the laws of nature are such that the early stages of the bus process cannot lead to Edna's death (for example, a world where buses and people repel each other, so that Edna and the bus cannot make contact with each other).

Staying with Dowe's general strategy, we can say that when relativizing the chance of e to the brick process we go to the closest world (of whichever kind) in which the bus cannot knock Edna over. For the purposes of the bus–brick case, it does not matter which kind of world (or which world of a particular kind) we go to, since in all three kinds of world the push does not hinder Edna's death, and hence – relative to the brick process – c increases the chance of e. (Note, however, that in none of the possible worlds described above does Dowe's bus-avoiding 'process', involving negative events $\neg b$ and $\neg d$, fail to occur – since b and d do not occur in any of them.)

The thought that the bus–brick case is in *some* sense a case of 'mixed paths' thus still seems plausible. So perhaps we could hang on to the chance-relativizing strategy in general without adopting Dowe's analysis of the 'process' from which we need to abstract away. The basic idea would go something like this: in cases where c lowers the chance of e but nonetheless causes e, e would have some (higher) chance of occurring had *c* not occurred. So there must have been some way for *e* to come about without the help of *c*: there must be some possible causal process – one that does not run to completion in the actual world because of the interference of *c* – which, if it ran to completion, would cause *e*. In the bus–brick case, this process is the process involving the bus speeding along the road, hitting Edna, and so on. It is the presence of the early stages of this process (plus surrounding circumstances and the laws of nature) in the actual world that makes *c* lower rather than raise *e*'s chance, since, were there no such early stages of such a process, or were there a concrete wall in the way, or were the laws different, *e* would have no way of occurring in the absence of *c*, and *c* would automatically raise *e*'s chance (from zero to something bigger).

In Dowe's terminology, it's the presence of the earlier stages of this potential causal process (plus surrounding circumstances and the laws) that makes *c hinder e*: by interrupting the more reliable way of producing *e*, *c*'s occurrence has the potential to prevent (that is, successfully hinder) *e*. So, if we want to restore chance increase between *c* and *e*, we need to think of a situation where that alternative potential causal process either does not get off the ground at all, or cannot run to completion. In such a situation, *c* will *not* be a hinderer of *e* – not because *c* somehow fails to initiate a hindering 'process' of negative events, but simply because *c* no longer has the capacity to prevent *e* because, without *c*, *e* would not occur.

Unfortunately, however, additional problems for this strategy arise when we try to accommodate chance-lowering causes that are not obviously analogous to the bus–brick case. Consider two other cases of chance-lowering causation to which Dowe applies the chance-relativizing strategy: the pulled-drive case described earlier, and the case of the decaying atom. In the pulled-drive case, Sue pulls her drive (*c*), thereby lowering the chance of the ball landing in the hole (*e*): had Sue not pulled the drive but instead struck the ball as she had intended, the chance of *e* would have been higher. In the decaying-atom case, an atom can decay to state *k* (call this event *e*) via either of two paths, one involving the intermediate product *i* and one involving intermediate product *j*. Either way – whether the process runs via *i* or via *j* – there is, at the time that *i* or *j* occurs, some chance that *e* will occur; but *j* gives *e* a higher chance than *i* does. Also suppose that time is discrete, and that the relevant steps are right next to each other: there are no relevant events in between the above-mentioned steps in the process, nor indeed any times for any such events to occur at. In fact the atom decays via *i* (let *c* be the event of its decaying to *i*); but prior to doing so it could have decayed via *j* instead. *c* therefore decreases the chance of *e* but, according to Dowe, *c* causes *e*.

While both of the above cases are similar to the bus–brick case in that *c* is (allegedly) a cause of *e* yet lowers *e*'s chance, they are disanalogous to it in that there is only one genuine causal process going on, rather than two. In each case, there is (intuitively) a single causal process going on, of which *c* is a part; and *c* modifies that process in such a way as to lower *e*'s chance. In the bus–brick case, by contrast, there were two separate processes – one involving the bus and one involving the brick.

How are we to think of the ‘mixed paths’ in the above cases? Well, in each case (as with the bus–brick case) *c* not only initiates a genuine causal process (the atom decaying to state *k*; the pulled drive, through the trajectory of the ball to the hole-in-one), but also acts as a hinderer (in Dowe’s sense) of *e*, since *c* has the capacity to prevent *e*: had the pulled drive resulted in what it was most likely to result in, it would have prevented a hole-in-one, for example.

Above I argued that the correct way to restore chance increase in the bus–brick case was to go *either* to the closest world where the earlier stages of the process that could, in *c*’s absence, have led to *e* do not occur, *or* to the closest world where those earlier stages occur but are somehow incapable of leading to *e*. So the crucial question is whether we can abstract away from that alternative process in the pulled-drive and decaying-atom cases too. I shall argue that *if* it is possible to do this in these cases, it is possible to do it in *all* cases of chance decreasers that are linked by a causal process to their effects. This yields the result that *all* such chance decreasers are causes – for example, on Dowe’s analysis the defoliant causes the plant’s survival.

Now, we saw earlier with regard to the bus–brick case that there were two possible ways of abstracting from the alternative process – that is, of going to a world where *c* is not a potential preventer of *e*. First, we could go to a world where the actual, earlier stages of the alternative process do not occur (a world where there is no bus at all, for example). Or, second, we could go to a world where those earlier stages *do* occur, but, for some reason, do not have the capacity to cause *e* (a world with different laws of nature that somehow do not permit the speeding bus to hit Edna, or a world where there is a wall between Edna and the bus, for example).

However, we do not have such a choice in the decaying-atom and pulled-drive cases: we have to use the second option. This is because earlier, actual stages of the alternative process that has the capacity to cause *e* are *also* earlier stages of the process which in fact, via *c*, leads to *e*. In the pulled-drive case, earlier actual stages of the alternative process – the process that could, without *c*, have led to *e* – include, for example, Sue putting the golf ball on the tee, lining up for the drive, and taking a swing. But those events are also part of the actual causal process that in fact led, via *c*, to *e*. If we go to a world where *those* events do not occur, we’ll find that *c* does not occur either: one cannot pull one’s drive without there being a ball to drive or a drive to pull. Similarly, for the decaying-atom case, the alternative process that might have led to *e* had *c* not occurred has as its earlier stages the continued existence of atom *h* just prior to the time when *c* occurred; *c* prevents that process running, via *d*, to completion. If we go to a world where those actual earlier stages are not present, we go to a world where there is no atom at all, and hence to a world where *c* cannot occur. Hence worlds where (actual) earlier stages of the alternative process do not occur will not be worlds where *c* increases the chance of *e*, since they will be worlds where *c* does not occur at all.

The reason why we cannot avail ourselves of the first option in these two cases is, of course, the fact that – unlike the bus–brick case – there is only one genuine causal process going on. The earlier stages of the causal process that in fact leads to *e* consist of the very same events as the earlier stages of the causal process that might, in *c*'s absence, have led to *e*. We therefore need to use the second option and go to a world where those earlier stages occur but do not, for some reason, constitute the earlier stages of a process that could, in *c*'s absence, cause *e*.

How are we to do this? Well, in the decaying-atom case, we have to go to a world where there is only one possible decay path to *k* – namely the path via *i*. At such a world – as Dowe points out (2000a: 80) – *c* really does raise the chance of *e*, since, at that world, were *c* not to occur, *e* would, by stipulation, have no chance of occurring. In the pulled-drive case, we have to go to a world where, for some reason, Sue simply has no way of getting a hole-in-one if she doesn't pull her drive. (Imagine, for example, that there is an obstacle that blocks the ball if it's on a straight-drive trajectory to the hole, but not if it's on a pulled-drive trajectory. Or imagine a world where the laws of nature are such that a 'straight' drive would produce a trajectory that would take the ball nowhere near the hole – or a conveniently located tree.)

This sort of world is, I think, the kind of world Dowe has in mind when he talks about going to a world where the hindering process does not occur – though, as I argued earlier, it is not the sort of world where, on his *analysis* of hindering 'processes', the hindering process does not occur. And this strategy really does, as Dowe claims, save (IC) and therewith (CCT), since the worlds described above are indeed worlds where *c* increases the chance of *e*. So the chance-relativizing strategy really does restore chance increase in cases of causation where *c* lowers the all-things-considered chance of *e*.

It seems, then, that although Dowe's own analysis of mixed-path cases (according to which the hindering 'process' is construed as a non-causal process involving negative events) does not succeed, his general strategy of abstracting away from the alternative causal process – the one in virtue of which *c* counts as a hinderer of *e* – successfully restores (relativized) chance increase in cases of (all-things-considered) chance-decreasing causation. However, a successful defence of (IC) that gives credence to common-sense intuitions about chance-decreasing causes must yield the result that some, *but not all*, chance decreasers that are linked to their effects by a genuine causal process are causes of those effects. Recall the defoliant example mentioned earlier: a plant is sprayed with defoliant (*c*) yet, six months later (after a period of being rather sickly), is in perfect health again (*e*). *c* lowers the chance of *e*, and there is a causal process between *c* and *e*. Philosophers who have discussed the case all agree (in a rare case of consensus over a thought experiment concerning causation) that *c* did not cause *e*.

Unfortunately, the strategy of abstracting away from the alternative causal process in order to restore chance increase can be applied just as easily to such cases of chance-decreasing causal processes as it can to cases of alleged chance-decreasing causation. For example, the defoliant case is susceptible to just the sort of move made above with respect to the decaying-atom and pulled-drive cases. The plant would have had a (higher) chance of survival had it not been sprayed (*c*) because its normally functioning processes would have been very likely, in the absence of *c*, to continue to operate and to lead to its survival in the usual way. But we can abstract away from that process, just as we did in the other cases. That is, we can go to a world where, for some reason, failure to be sprayed would lead to the death rather than the survival of the plant. For instance, we can imagine a world where, without the spraying, the plant would succumb to a disease that can only be transmitted through its leaves. At that world, *c* increases the chance of *e*, since without *c* the plant would not survive. So, relativized to the actual causal process between the spraying and the survival, *c* increases the chance of *e* – so on Dowe's account, the spraying caused the survival.

It is clear that the chance-relativizing strategy will generalize to all cases where there is a causal process between *c* and *e*. In all cases of causal process plus chance decrease, there must be some alternative potential route that *c* prevents or cuts short, which might, in the absence of *c*, have led to *e*. If there were no such alternative potential route to *e*, *c* could not lower the chance of *e* in the first place, since without *c* there would be no chance of *e* occurring at all. Once we see this, and see that the chance-relativizing strategy in effect abstracts away from the possibility of that alternative causal process running to completion, it is easy, for every chance-decreasing causal process, to cook up a possible world where that alternative process is unable to run to completion.

Dowe's attempt to rescue (IC) therefore fails, since it has the consequence that the existence of a causal process between *c* and *e* is a sufficient condition for causation: the attempt renders not just some, but *all* cases of chance-decreasing causal processes as cases of causation.

3 Causing, hindering and the bullet-biting strategy

In my article 'Taking Hindrance Seriously' (1997), I present an analysis of causation according to which hindrance is a kind of causal relation. Ignoring the question of how a causal process is to be defined – an issue that is not directly relevant to the purposes of this paper – the analysis's central principle is the following:

(H) *c* and *e* are causally related if and only if there is a causal process between them. If there is a causal process between *c* and *e*, then *c* causes *e* if and only if *c* increases the chance of *e*, and *c* hinders *e* if and only if *c* decreases the chance of *e*.[10]

(H) entails that there is no such thing as chance-decreasing causation, since it entails that alleged cases of chance-decreasing causation, like the pulled-drive case, the bus–brick case and the decaying-atom case, are in fact cases of hindrance rather than causation. My strategy for rescuing (IC) is therefore the rather simplistic strategy of biting the bullet.

In this section I argue that, if we accept (H), the bullet-biting strategy is a perfectly reasonable strategy, since (H) provides the resources to ease the pain somewhat. First, however, I show that analyses of indeterministic causation that do not take hindrance to be a species of causal relation face a serious problem that does not arise for (H), and that such analyses do justice to intuitions about alleged chance-lowering causation at the expense of violating intuitions about chance-lowering causal processes that are *not* cases of causation – the defoliant case, for example. I also offer a somewhat speculative diagnosis of why standard analyses of causation fail to accord hindrance the status of a causal relation and argue that the reasons for failing to do so are bad reasons. In the next section, Section 4, I argue that once (H) is accepted, some objections raised by Dowe to the bullet-biting strategy can be answered.

I argued in Section 2 that Dowe's attempt to restore chance increase to cases of chance-decreasing causation fails because it makes *all* chance-decreasing causal processes come out as cases of causation. In fact, it is not very surprising that Dowe's analysis should have as a consequence that all causal processes – whether chance-increasing or chance-decreasing – turn out to be cases of causation, since for Dowe hindering is not a species of causal relation: hindering 'processes' are not *causal* processes. Roughly speaking, for Dowe *c* hinders *e* if and only if *c* cuts off some process which would, had *c* not occurred, have been more likely to cause *e*. So *c* hinders *e* not in virtue of the *actual* causal process that runs from *c* to *e*, but in virtue of the fact that *c stops* some *other* causal process from running to completion. Since there is, by stipulation, a causal process between *c* and *e*, and since the fact that *c* hinders *e* obtains not in virtue of the existence of *that* causal process but rather in virtue of the non-existence of some other causal process, when we ask what the *causal* relation is between *c* and *e* – the relation that obtains in virtue of the existence of a causal process between *c* and *e* – the only available answer is that *c* causes *e*.

Other analyses similarly have no room for hindrance as a species of causal relation. Lewis (1986), for example, defines causal dependence as chance increase, and then defines causation as a chain of causal dependence. This has the effect of 'factoring out' any chance decrease: when *c* lowers the chance of *e* there is generally some chain of events (*d* and *f*, say) such that *c* increases the chance of *d*, *d* increases the chance of *f*, and *f* increases the chance of *e*. In all such cases – the defoliant case, for example – *c* comes out as a cause of *e*. Menzies's (1989) analysis produces the same result. Dowe calls such accounts 'interpolating' accounts because they attempt to hook a chance-decreasing cause to its effect by interpolating further chance-increasing events between cause and effect.

Interpolating accounts in effect render all causal processes as cases of causation; like Dowe's 'mixed paths' analysis, they turn the causal process condition into a sufficient condition for causation. Hence they only yield the 'right' result in the pulled-drive and bus–brick cases at the expense of making all chance-decreasing causal processes come out as cases of causation. So they save intuitions about chance-decreasing causes at the expense of violating intuitions about chance-decreasing non-causes – for example the intuition that the spraying was not a cause of the plant's survival.

Moreover, interpolating accounts run into worse trouble in alleged cases of 'direct' causal relations. In Dowe's decaying-atom case, for instance, we are supposed to imagine that time is discrete, so that no further events can be interpolated between *c* and *e* in such a way as to yield a chain of chance-increasing causal dependence. Since interpolating accounts recognize only one kind of causal relation – causation – they do not have the resources to recognize any kind of causal relation at all between *c* and *e* in the decaying-atom case, rendering *c* and *e* completely causally unrelated. Recognizing hindrance as a species of causal relation, however, provides a simple solution to the problem posed by the decaying-atom case, since (given an appropriate characterization of a causal process) (H) allows us to say that *c* and *e are* causally related, since *c* hinders *e*.[11]

Why is it that analyses of causation typically recognize only one kind of causal relation – namely causation? Well, one way of explaining it is to see it as a hang-over from deterministic analyses of causation. Under the assumption of determinism, there really is no hindrance (construed as a chance-decreasing kind of causal relation), since to lower *e*'s chance under determinism is to lower it from 1 to 0; hence it is impossible for *c* to decrease the chance of *e* and for *e* still to occur. Once we abandon determinism, however, there is no good reason for thinking that genuine hindrance does not exist, or that it is somehow not a genuinely causal relation. With indeterminism in place, there is no reason to regard chance decrease as any less real or any less important than chance increase when it comes to analysing causation, and hence no reason to regard the kind of causal relation manifested by chance-decreasing causal processes – hindrance – as any less real or important than the kind – causation – manifested by chance-increasing processes.

One might object that the term 'hinders' can perfectly appropriately be used in deterministic contexts, and hence that hindrance cannot (as I've claimed) be a feature of the world that only exists if indeterminism is true. For example, suppose I succeed in lighting a damp match. Whether or not the relevant processes are deterministic, it seems appropriate to say that the match's dampness hindered its lighting. But on my account of hindrance, in the deterministic case the dampness didn't *really* hinder the lighting. When I struck the match, circumstances were such as to guarantee that the match would light. The dampness of the match may or may not have been a necessary condition of the match's lighting. If it was, then the dampness was in fact a cause of the lighting, and if it wasn't, then the dampness was simply irrelevant to the lighting. Hence (so the objection goes) whatever the relation is that I've called 'hindrance', it cannot really *be* hindrance, since intuition tells us that the term 'hindered' applies in cases – namely, deterministic cases – where on my account the term does not apply.

This objection is not a serious one since, so far as I can tell, the desire to say that the dampness hindered the lighting simply goes away if one takes the deterministic starting point of the objection seriously. If the dampness was necessary for the lighting (say because I only struck with sufficient force because I knew the match was damp – if it hadn't been damp, I would not have taken such care, and it would not have lit), then it seems perfectly appropriate to say that the dampness caused, rather than hindered, the lighting. And if the dampness made no difference to whether or not the match lit, it seems perfectly appropriate to say that the dampness was irrelevant to the lighting.

In fact, my use of the term ‘hindered’ does not differ so very much from the deterministic usage described above. According to (H), hindrance is an indeterministic relation characterized by chance decrease. Deterministic usage of the expression ‘hindered’ can be seen as expressing the idea that *c* lowers the *probability* of *e*, where the probability of *e* is a probability of the non-single-case variety – the kind that can take values other than 0 and 1 even if determinism is true. Generally speaking, damp matches light far less frequently than dry ones; hence, given some incomplete specification of the circumstances in which I strike the match, the dampness lowers the probability (though not the chance) that the match will light. Of course, the non-single-case probability of *e* will vary depending on how the situation is described, and hence there is no univocal answer to the question ‘did *c* hinder *e*?’ in deterministic situations. This is not the case with hindrance as defined by (H), since single-case chances are not relative to how the relevant events are described.

We are still left with the problem set up at the beginning of the paper – that of chance-decreasing causation – since (H) is incompatible with its existence. According to (H), Sue’s pulled drive hindered (and did not cause) the hole-in-one, the atom’s decaying to state *i* hindered (and did not cause) its decaying to state *k*; and the push hindered (and did not cause) Edna’s untimely death.

According to Dowe (and no doubt according to others too), such results go against our intuitions: intuitively, in all three cases, *c* caused *e*. I do not think the case for respecting intuitions is particularly strong here. Recall that on standard analyses of causation there is no distinction between ‘*c* and *e* are causally related’ (or ‘there is a causal process between *c* and *e*’) and ‘*c* causes *e*’. So, on such analyses, to deny that *c* caused *e* is to deny that there is any kind of causal relation or connection whatsoever between *c* and *e*. Given this starting point, the desire to say that *c* caused *e* in the above cases is natural, since in those cases it would indeed be implausible to claim that *c* bears no causal relation to *e* at all. However, once we hold that causation is not the only kind of causal relation, we can safely deny that *c* caused *e* without rendering *c* and *e* causally unrelated. And this is precisely what taking hindrance to be a species of causal relation allows us to do: to say that *c* hindered *e* is to assert that *c* and *e* are causally related; but it is also to deny that the causal relation thereby instantiated is the relation of causation.

One might nonetheless claim that brute intuitions in the three cases of alleged chance-lowering causation should be respected. Perhaps if there were some viable analysis of causation that yielded the result that *c* caused *e* in all three cases, that would be a good reason to prefer that analysis. However, as I have argued, the prospects for such an analysis are dim.

4 Dowe’s objections to the bullet-biting strategy

Dowe raises two objections against the bullet-biting strategy of denying that alleged chance-decreasing causes really are causes. First, he points out that the bullet-biting strategy entails that causation is not transitive, since according to that strategy it can be the case that *c* causes *d*, *d* causes *e*, but *c* hinders (rather than causes) *e*. The bus–brick case is an example of this: the push causes Edna to be on the pavement, which in turn causes her to be hit by the brick, which in turn causes her death. But the push hinders, rather than causes, the death.

It is certainly true that most philosophers are very keen to hang on to the transitivity of causation. Douglas Ehring, for example, claims that ‘transitivity is a fundamental logical feature of the causal relation … causal transitivity should be disavowed only as a last resort’ (Ehring 1997: 82). However, I have not been able to find any *arguments* for the thesis.[12] Perhaps one reason why transitivity is so popular is that it functions as a kind of methodological principle. When we want to trace the causes of some event – Edna’s death (*e*), say – we typically start by identifying the event’s immediate causes (being hit by the brick (*d*), say), and then tracing back to the causes of *those* events, and so on. It seems that if causation were not transitive, there would be no guarantee that such a procedure would reveal distant causes of Edna’s death, since without transitivity the fact that each step was caused by the preceding step is no guarantee that the first step caused the last.

Does denying transitivity amount to denying the general applicability of this valuable methodological principle? Well, no – not on the view proposed here. On my view, the ‘there is a causal process between’ relation *is* transitive. So, in tracing the steps in the chain of events that led to Edna’s death, we are identifying the *causal process* that led to it. For example, when we establish that the push (*c*) caused *d*, and that *d* caused *e*, we have – by the transitivity of causal processes – established that there is a causal process between *c* and *e*. The only difference is that, before jumping to the conclusion that *c caused e*, we have to determine whether or not *c* raised the chance of *e*. In the current case, it turns out that *c* did not raise *e*’s chance; hence we should conclude that *c* hindered *e*.

It’s worth reiterating the point that hindrance (so defined) is a phenomenon that can only occur if the world is indeterministic. In deterministic situations, the existence of a causal process between *c* and *e* guarantees that *c* causes *e*, since it is impossible for there to be a causal process between *c* and *e* if *c* lowers the chance of *e* from 1 to 0. Given that many analyses of causation presuppose determinism, or are descendants of analyses that presuppose determinism, it is therefore not surprising that it should be so widely taken for granted that causation is transitive – since, assuming determinism, there is a causal process between *c* and *e* if and only if *c* causes *e*.[13] It is only when we abandon determinism that causation and causal processes come apart and we can legitimately ask whether one of the two relations might not be transitive. And I can see no reason not to deny the transitivity of causation so long as we hold on to the transitivity of causal processes. Indeed, giving up on the transitivity of causation is the only way of getting the right result in the defoliant case. There is clearly a causal process running from the spraying to the survival – which is to say, there is a chain of events running from the spraying to the survival such that each event in the chain caused the next. So the only way to deny that the spraying caused the survival is to deny that the existence of that chain of causation is sufficient for the first step to cause the last.

Dowe's other objection is that if we deny that alleged cases of chance-decreasing causation really *are* cases of causation, we make the obtaining of the causal relation depend on extrinsic features of the situation; and, intuitively, causation is an intrinsic matter. As Peter Menzies puts it:

> I drop a piece of sodium into a beaker of acid and that causes an explosion, this causal relation is an intrinsic feature of the cause-effect pair. So if there is another person waiting in the wings, ready to drop a piece of sodium into the beaker if I do not, that makes no difference to whether the causal relation holds between my dropping the sodium and the explosion. The presence of the alternative cause is neither here nor there to the causal relation that exists between the actual cause and effect. The causal relation does not depend on any other events occurring in the neighbourhood: the causal relation is intrinsic, in some sense, to its relata and the process connecting them.
>
> (Menzies 1998: 339)

Dowe's objection is that denying that, for example, the bus–brick case is a case of causation amounts to denying that the causal relation does not depend on any other events occurring in the neighbourhood, and therefore makes causation unacceptably extrinsic. He says:

> it seems implausible to suppose that whether the push caused the death is dependent on how fast a bus was going, a bus what's more that didn't hit her. Suppose the chance that the lady dies given the push is 0.5. Suppose if the bus is travelling at 36 miles per hour that the chance that the old lady will die given that there is no push is 0.45, and that if the bus is travelling at 38 miles per hour that the chance that the old lady will die given that there is no push is 0.55. Then, if the bus is travelling at 36 miles per hour when the push occurs, which leads to her being hit on the head by a brick and dying, then the push caused her death. But if the bus is travelling at 38 miles per hour when the push occurs, which leads to her being hit on the head by a brick and dying, then the push does not cause her death. The intuition is that whether an event causes another via a particular process shouldn't depend on the strength of a separate, distant unsuccessful process.
>
> (Dowe 2000a: 75)

The issue of whether causation is an intrinsic relation is a difficult and controversial one (not least because one's answer depends on how one defines 'intrinsic'[14]). I shall, however, ignore the technicalities of this debate and assume some sort of pre-theoretic understanding of 'intrinsic' and 'extrinsic', according to which, for example, how fast the bus is going is clearly extrinsic to the causal process initiated by the push, whereas features like the strength of the push and the velocity of the brick are intrinsic features of the process. I shall argue that both Menzies and Dowe are here running together what are, on my view, two separate questions: the question of whether the *causal process between* relation is intrinsic, and the question of whether the *causal* relation is intrinsic.

Before doing that, however, let's look at the broader picture for a moment. There is a general tension between, on the one hand, thinking of causation as an intrinsic relation and, on the other, holding (as most contemporary theorists do) that an adequate analysis of causation must take account of the chances that causes give their effects. (This is meant to include not just theories that appeal directly to chance in their analysis of causation, but also to theories that simply seek to do justice to the thought that causation and chance are somehow conceptually connected. For example, a commitment to (CCT) displays a commitment to a connection between causation and chance, whether or not (CCT) actually plays a part in one's analysis of causation.)

This tension arises because whether or not *c* raises the chance of *e* depends not just on the intrinsic features of *c* – and not even just on the intrinsic features of *c*, plus the intrinsic features of the causal process running from *c* to *e*, plus the laws of nature – but on facts that are entirely extrinsic to *c* and to the causal process that thereby leads to *e*. In the bus–brick case, for example – as Dowe notes – *c* lowers rather than raises the chance of *e* because of the speed at which the bus happens to be going. In the defoliant case, *c* lowers rather than raises the chance of *e* because in fact external factors are not such as to jeopardize seriously the plant's chances of survival were all its leaves to remain intact. Whether or not *c* raises the chance of *e* depends in part on what the chance of *e* would have been in the absence of *c*; and that generally depends on features of the world that are nothing to do with the intrinsic features of the actual causal process between *c* and *e*.

Moreover, whether *c* raises or lowers the chance of *e* also depends on what the *actual* chance of *e* is; and this also often depends on facts that are extrinsic to the causal process linking *c* and *e*. Suppose, for example, that the defoliant only has a 50% chance of affecting the plant's leaves at all. Then the plant's *actual* chance of survival just after it is sprayed depends not only on how likely it is that the plant will die if it loses its leaves, but also on how likely it is that it will die if its leaves remain intact. Suppose that circumstances are such that if the plant keeps its leaves (whether or not it is sprayed) it will have a 90% chance of survival, but that if it loses its leaves (which it can only do if sprayed) its chance of survival will be just 20%. Then just after spraying, the plant's actual chance of survival is 55%. Now suppose that circumstances are such that if it keeps its leaves it will have only a 10% chance of survival (because of the presence in its immediate environment of a nasty leaf disease, say). Then just after spraying, the plant's actual chance of survival is 15%. Thus extrinsic facts – facts about the surrounding environment – help determine the actual chance of *e*, and thus whether or not *c* raises the chance of *e*, even though those extrinsic facts play no part in the actual causal process leading from the spraying to the survival. It is therefore very difficult to hold on to the idea that causation is an intrinsic matter while at the same time holding that causation and chance are related.

On the view of causation as a chance-increasing causal process, causation is (at least partially) extrinsic in a very obvious way, since, as we saw above, whether or not *c* increases the chance of *e* is at least partially an extrinsic matter. Is this a counterintuitive result? I think not. For one thing, intuitions about whether *c* causes *e* sometimes *do* hinge on extrinsic features of the situation. So to rule that causation must *never* depend on extrinsic features would violate those intuitions. Consider the defoliant case. As the case is actually set up (without the story about the leaf-infecting disease), common-sense intuition favours the verdict that the spraying was not a cause of the plant's survival. But now consider the variant case where, in the absence of the spraying, the plant would be highly susceptible to a fatal disease. In *that* case, I think common-sense intuition favours the verdict that the spraying *was* a cause of the plant's survival. Yet the differences between the original case and the variant case are purely extrinsic differences: the intrinsic nature of the actual causal process leading from the spraying to the survival are precisely the same in each case.

Second, given (H), we must remember that on the account offered here, the question of whether *c* causes *e* must be distinguished from the question of whether there is a causal process running from *c* to *e*. With the distinction in place, we can accept that causation itself is (at least partially) extrinsic without being required to accept that the existence of a causal process between *c* and *e* is likewise an extrinsic matter. That is, the extrinsicality of causation is perfectly compatible with a fully intrinsic characterization of the notion of a causal process.[15] Granted, on my view whether or not the push is a cause of Edna's death depends in part on how fast the bus is going. But whether there is a causal *process* between the push and the death does not depend on how fast the bus is going, nor even on whether there is a bus on the scene at all.

Both the transitivity objection and the intrinsicality objection can be met, then, by separating questions about causation from questions about causal processes, and showing that according to (H), the causal-process-between relation can meet the desiderata even though causation itself does not. For the objections to stick, one would have to show that, once the two relations have come apart, it is causation, rather than (or perhaps as well as) the causal-process-between relation to which transitivity and intrinsicality properly apply. I can see no reason to think that this can be done. Once (H) is accepted, then, the bullet-biting strategy is a plausible strategy for saving (IC): we can retain the view that causes increase the chances of effects without appealing to the chance-relativizing strategy.

Notes

1 For attempts at a formal definition of a causal process, see for instance Dowe (1992) and Beebee (1997).

2 This latter example is perhaps more controversial than the others are. One might, for instance, hold that there is a causal process of negative events – including, say, the emptiness of the fridge, my not chopping onions, and so on – linking my failure to go shopping to my failure to cook dinner (provided that an appropriate relation – counterfactual dependence, say – holds between each step). Anyone who holds the view that causal processes can obtain between absences, or between events and absences or

absences and events, should note that I mean by 'causal process' the more common-or-garden kind – a kind that does not count negative events as participants in genuine causal processes.

3 The defoliant case is due to Nancy Cartwright (1979).

4 Assuming, of course, that we could provide satisfactory analyses of causal processes and counterfactual increase in chance – no easy task, and not one I shall attempt here.

5 I make a start on this job in my article 'Causing and Nothingness' (forthcoming), where I defend the view that there is no causation by absence.

6 Dowe claims that the pulled drive is a cause of the hole-in-one, though others (including Hugh Mellor, whose example it is) disagree; see Mellor (1995: 67–8).

7 See Dowe (2000a: 71) for this example, henceforth called 'the bus–brick case'.

8 See Dowe (2000a: 69) (I have changed the wording slightly).

9 Of course, in some intuitive sense the σ-process in the actual world and the σ-process in the bus-free world look very different from one another, since *b* and *d* fail to happen in virtue of very different kinds of positive events. Still, assuming negative events occur by definition just if the positive ones don't, they really are the same 'events' at each world, and hence the same 'process'.

10 As stated, the analysis says nothing about the case where there is a causal process between *c* and *e* but *c* does not change the chance of *e* at all, and thus remains silent for cases of deterministic preemption where *c* leaves the chance of *e* at 1 – just what it would have been in the absence of *c*. This omission can be rectified by stipulating that *c* causes *e* if and only if there is a causal process between *c* and *e* and *c* does not decrease the chance of *e*.

11 I present the argument that interpolating accounts cannot deal with direct chance-decreasing causal relations in more detail in my (1997).

12 In a recent paper, E. J. Hall (2000) also claims that transitivity is too important to give up – and that we should only do so if we have independent reason for it. However, Hall is concerned solely with deterministic causation, where there is no distinction between the existence of a causal process and causation. I think the move to indeterminism *does* provide an independent reason to give up the transitivity of causation – but not, as we shall see, the transitivity of causal processes.

13 Recall that I am presupposing that (CP) is true.

14 See for example Langton and Lewis (1998).

15 In fact, on my own view (indeed on most views), causal processes are not wholly intrinsic either. But they are at least *reasonably* intrinsic in the sense that whether there is a causal process between, say, the dropping of the sodium and the explosion does not depend on whether there is someone waiting in the wings to drop the sodium in if I do not; and whether or not there is a causal process between the push and Edna's death does not depend on how fast the bus is going.

Chapter 5

Counterfactual theories, preemption and persistence

Douglas Ehring

The claim that causation is ultimately reducible in some way to some form of counterfactual dependence might appear to be a nonstarter. We are all familiar with causes accompanied by failsafe mechanisms. A back-up shooter, ready and waiting in case Oswald failed, would have made Oswald unnecessary, but not ineffective. Not surprisingly, counterfactual theorists recognized preemption from the beginning (Lewis 1986) as a primary problem case and still do. As Lewis said as recently as 2000, the simplest counterfactual account breaks down in cases of redundant causation, including preemption, 'wherefore we need extra bells and whistles'. And, indeed, there has been no shortage of preemption-based bell-and-whistle development. In this paper, my goal is to show that a number of the leading variants of the counterfactual theory designed in part in response to preemption cases cannot handle a certain form of preemption (which might be called 'persistence preemption'). A larger goal is to bolster a line of argument from my *Causation and Persistence* (1997) in which I argued that a theory of causation, even a generalist one, should include a singularist component involving persistence.[1] The first stage of that argument was comparative. Transference theory (read as singularist, for purposes of that argument) was shown to do a better job with this kind of preemption than other reductionist theories, including counterfactual theory. I then discarded transference theory to develop an alternative singularist component of a generalist theory of causation, arguing for a trope persistence interpretation of that component. Here I reinforce the comparative part of that argument, extending it to counterfactual accounts not examined earlier. To that end, I introduce a variant of a kind of preemption case I considered earlier (Ehring 1997: 42) and then show that various counterfactual theories cannot handle this case, but theories with a singularist-persistence component can. Again I use a naïve form of transference theory as an example of a theory of the latter sort.

Modifications of the simplest form of counterfactual theory have tended, with some exceptions, to focus on one and/or the other of two aspects of many preemption cases: (1) had the preempting cause not occurred and the preempted cause given rise to the effect, the causal chain leading to the effect would have included some events that had not in fact occurred; and (2) had the preempting cause not occurred and the preempted cause given rise to the effect, the effect would have occurred later or been different in some relevant respect. Preemption cases that lack these features are unlikely to be compatible with modified versions of counterfactual theory that focus on these contingent features of some preemption cases. I described such a case (Ehring 1997) involving the collision of particles. I modify that case here and recast it to fit the 'node' framework common in many discussions of counterfactual theory.

Case A

There are two qualitatively indistinguishable particles with equal quantities of energy moving on a collision course. The laws mandate that if these two particles collide, one and only one particle will be destroyed (each has a 50% chance of annihilation). For the particle that is not destroyed, there is a chance that it will jump (noncontinuously) from the point of collision across a spatial gap to a location *e* some distance away. The laws dictate that if one particle reaches the location of the collision without the other, then there is a chance that it will jump to *e*. Now as a matter of fact there is a collision and one, but not the other, particle jumps noncontinuously to *e* and the other particle is annihilated. The effect of interest is the presence of a particle of a certain type with a certain amount of energy at *e*, after the collision. Depending upon which particle jumped we have different causal stories. For example, if the first particle does the jumping then its presence at the collision point and its possession of energy at that point is the cause.

Now let's recast this example in the 'node' format (Figure 5.1). There are two series of nodes. A node's 'firing' at *t* consists of the presence at that node of a particle with a certain quantity of energy, *q*. Particles 'jump' from node to node without occupying the space between them. Particles are transmitted probabilistically from node to node. The 'firing' of a node increases the probability of the next node's firing.[2] Without interference, each link in the *a*–*e* chain is much more reliable than each corresponding link in the *b*–*e* chain.

Nodes *d* and *g* are adjacent to each other. If particles arrive at *d* and *g* at the same time because of their proximity, there is a collision. The result of a collision is the

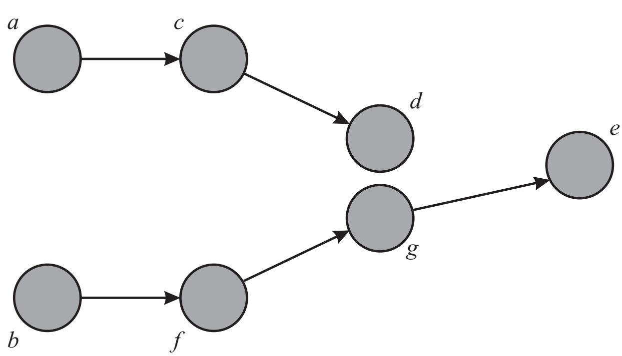

Figure 5.1

annihilation of one particle or the other, but not both. The collision acts as a two-way inhibitory mechanism, running in one direction or the other, but not both, with an equal chance of running in either direction. When a node is 'inhibited', it fires (the particle/energy is present) at the time it would have absent inhibition, but the particle/energy disappears in the next moment. If the node is inhibited at *t* by a collision, then it fires at *t* but it does not transmit a particle in the next unit of time, *t′*. The quantity of energy possessed by that node at *t* is destroyed along with the particle and does not exist at *t′*. If *d* is inhibited by *g* at *t*, then *g* fires at *t*, *d* fires at *t*, but *d*'s particle, and its energy, ceases to exist at the next unit of time, *t′*. Had sequence *a–c–d* or *b–f–g* occurred by itself, the firing of *d*/*g* would have raised the probability of *e*'s firing.

Now suppose on a particular occasion all nodes fire but the inhibitory mechanism/collision blocks the transmission of the particle/energy from *d* to *e*, and, instead, the particle/energy is transmitted to *e* from *g*. Hence, *e*'s firing is caused by *g*'s firing, not by *d*'s firing, and the *a–e* process is preempted by the *b–e* process. However, had *g*'s firing not occurred, the chance of *e*'s firing would have been higher given the greater reliability of *d* in transmitting particles/energy.

Schaffer (2000a) raises similar problem cases (his 'trumping preemption' cases) for counterfactual theories that should be distinguished from this case.

> *Schaffer's 'magic' case:*
> Merlin casts a spell earlier on the same day that Morgana casts a spell, both aimed at turning the prince into a frog at midnight. Merlin's spell preempts Morgana's spell in turning the prince into a frog since there is a law that the first spell cast on a given day match the enchantment that midnight.
> (Schaffer 2000a: 165)

What makes these cases similar is that in neither would the effect have occurred later nor would the effect have been brought about by some non-actual events had the preempted cause brought about the effect. Still these cases are importantly different. This difference can by brought out by reference to a criticism of Schaffer's 'trumping preemption' cases. According to McDermott (2002), contrary to Schaffer it is not the case that one spell process runs to completion, but the other does not. Processes that are non-causally intrinsically indistinguishable, in these circumstances, are not causally distinguishable: 'It is an intuition of "intrinsicness": the (one step) process from Morgana's spell to the prince's frogification runs to completion exactly as it would have done in the absence of Merlin's spell, and it certainly would have been a causal process in that case … . The so-called trumping case (with no intermediate events) is one of over-determination, not preemption.' (McDermott 2002: 89) What is important to note is that Schaffer's case is not consistent with the intrinsicness of causation but Case A is. This difference is evident in the fact that there is a response available for Case A against this line of criticism that is not available for the magic case.[3] The 'intrinsicness' objection cannot get a grip on Case A. There is a non-extrinsic

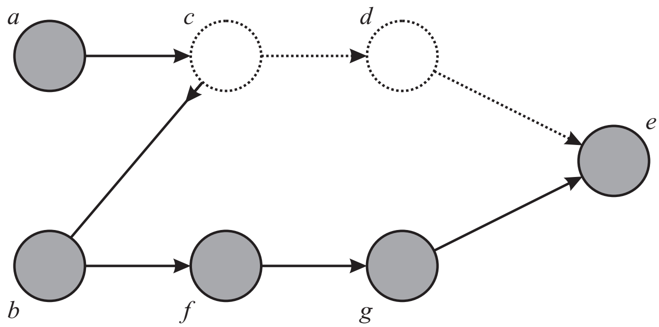

Figure 5.2

difference between the processes in Case A that grounds a causal difference. In one chain, but not the other, there is persistence of a particle/quantity of energy through to the effect node.

Now let's consider how variants of counterfactual theory might deal with Case A. I begin with Lewis's earliest treatment of preemption. As we shall see, it does not work on Case A.

Counterfactual theories

Lewis

Lewis's original formulation of the counterfactual theory requires only stepwise counterfactual dependence, not dependence (where if *c* and *e* are actual events, *e* depends on *c* just in case had *c* not occurred, *e* would not have occurred):

> For actual events *c* and *e*, *c* causes *e* just in case there is a series of actual events $x_1 \ldots x_n$ such that x_1 depends on *c*, x_2 on x_1, … , and *e* depends on x_n.
>
> (Lewis 1986: 167)

This leaves it open that preemption is consistent with counterfactual theory if preemptively caused effects always stepwise depend on their preempting causes. Initially, Lewis claimed just that about preemption. Let's illustrate his initial gloss on preemption with an example which is much like one from Lewis's work (1986: 200).[4] In Figure 5.2, circles are neurons. Filled-in circles indicate neuronal firings and empty circles are neurons that do not fire. Forward non-dashed arrows indicate stimulation and reverse arrows indicate inhibition. A neuron that is both stimulated and inhibited does not fire. All causation is deterministic.

b's firing causes *e*'s firing by way of the firings of *f* and *g*. *b*'s firing blocks *c*'s firing.[5] Lewis conjectured that there will be at least one intermediary (in this example, say, *f*'s firing) between the event which initiates the blocking action (here, *b*'s firing) and the effect. If *b*'s firing had not occurred, the firing of *f* would

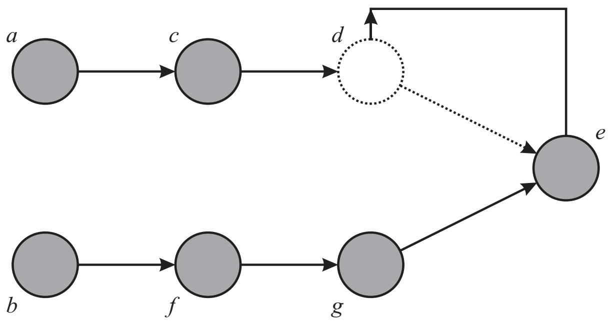

Figure 5.3

not have; and, given that by the time *f* fires, the alternative process is already doomed by the earlier firing of *b*, if the firing of *f* had not occurred, then both processes would still have failed to run to completion (Lewis 1986: 171, 200). Stepwise dependence is, thus, established. This approach does not work in Case A since it requires that the blocking-initiator event dooms the preempted process *before* occurrence of an intermediary event in the main line. In Case A, there is no intermediary event after the blocking-initiator event. The blocking-initiator event, *g*'s firing, is also the direct cause of *e*'s firing.[6]

In a later postscript to his original paper on causation, Lewis concluded that the 'stepwise' approach does not work for late preemption/late cutting. Consider the following case (Figure 5.3) of deterministic late cutting (Lewis 1986: 203–4).[7]

The preempted line is blocked (*d*'s firing is blocked) only by the effect event itself and nothing earlier. For every event in the preempting process, it is false that had that event not occurred, an *earlier blocking* event would still have occurred, dooming the preempted process. Assuming that the final effect could have had an alternate causal history and that it is not temporally fragile – that it could have occurred somewhat later – stepwise counterfactual dependence fails. Lewis revised his account as follows:

> *c* causes *e* just in case there is a series of actual events $x_1, \ldots, x_n$ such that x_1 depends or quasi-depends on c, x_2 on x_1, … , and *e* depends on or quasi-depends on x_n.
>
> (Lewis 1986: 206)

Quasi-dependence is characterized basically as follows:

> *e* quasi-depends on *c* just in case the intrinsic character of the process connecting *c* and *e* is just like that of processes in other regions of this or other worlds with the same laws and in the great majority of these regions these processes display stepwise counterfactual dependence.
>
> (Lewis 1986: 206)

Lewis conjectured that the preempting cause satisfied this condition in cases of deterministic late preemption (Lewis 1986: 206). Applied to the case in Figure 5.3, in the great majority of regions with processes intrinsically just like the one connecting *b*'s firing with *e*'s firing, there will be a chain of counterfactual dependence since preemptive settings are rare. Lewis also seems to presume that the same is not true of the preempted process. In the majority of other-regional processes with processes much like *a–e*, those processes will not be exactly like that of *a–e* since they will include an additional event in place of the missing firing of *d*, for which there is no analogue in the *a–e* process (*the missing event feature*).

In order to test the 'quasi-dependence' account against Case A, which is probabilistic, let's now bring in Lewis's probabilistic notion of dependence – required for indeterministic situations in which there is still some chance that the effect would have occurred spontaneously in the absence of the cause even without preemption/overdetermination (Lewis 1986: 176).

> For actual but distinct events *c* and *e*, *e* probabilistically depends on *c* just in case the actual chance of *e* occurring is *x* (where 'the actual chance *x* of *e* is to be its chance at the time immediately after *c*') and if *c* had not occurred, *e* would have had some chance *y* of occurring very much less than *x*.
>
> (Lewis 1986: 176–7)

The 'quasi-dependence' formulation, in effect, then becomes:

> For actual and distinct events *c* and *e*, *c* causes *e* just in case there is a series of actual events $x_1, \ldots, x_n$ such that x_1 probabilistically depends or quasi-probabilistically depends on *c*, x_2 on $x_1, \ldots$, and *e* probabilistically depends on or quasi-probabilistically depends on x_n.
>
> (Lewis 1986: 206)

The 'quasi-dependence' approach does not work for Case A. Consider processes in other regions that are just like the *a–c–d–e* process (Figure 5.1). In a majority of regions with the same laws in which there is realized a sequence which is event-for-event intrinsically indistinguishable from that preempted line, the final effect *does* stepwise probabilistically depend on events in that sequence. Those other-regional processes do not differ from the *a–c–d–e* sequence 'event-wise' since they do not include 'extra events'. Although a transmission is blocked in the preempted line, there is no cutting of events, no missing intermediary events.[8] One might say that there is blocking but no cutting. The preempted line satisfies the 'quasi-dependence' account.

Consider a possible reply:

> The failure of transmission/persistence in the preempted line is the relevant intrinsic difference between the preempted process and its other-regional

analogues in which there is no missing transmission. Missing transmission/persistence takes over for the missing events.

I have two responses to this defence. First, this defence will not be acceptable to Lewis. If the persistence of the particle or quantity of energy is itself to be analysed as a matter of causally connected temporal parts of that particle/quantity, that will be inconsistent with Lewis's *reductionist* counterfactual theory of causation. Or, if the persistence of this particle/quantity is analysed as a matter of that particle/quantity being wholly present at each moment that it exists, that is not consistent with Lewis *qua* proponent of temporal parts. Second, once we appeal to a persisting particle/quantity of energy in a defence of the 'quasi-dependence' view, that opens up the possibility of a theory of causation that includes a persistence component which appeals directly to the persistence-based difference between the preempted and preempting process instead of getting at this difference indirectly by way of relations of resemblance across regions.

Lewis (2000) proposes a new, 'influence' version of counterfactual theory: *c* causes *e* if and only if *c* and *e* are actual distinct events linked by a chain of causal influence. Causal influence is a matter of a mapping of counterfactual alterations of effects on to the counterfactual alterations of their causes. Lewis puts it as follows: 'Where *C* and *E* are distinct actual events, … *C influences E* if and only if there is a substantial range C_1, C_2 … of different not-too-distant alterations of *C* … and there is a range E_1, E_2 … of alterations of *E*, at least some of which differ, such that if C_1 had occurred, E_1 would have occurred, and if C_2 had occurred, E_2 would have occurred, and so on' (Lewis 2000: 190). An alteration of an event *e* is an event similar to *e* that need not be a version of *e*. Lewis says that an '*alteration* of *E*' is 'either a very fragile version of *E* or else a very fragile alternative event that is similar to *E*, but numerically different from *E*' (Lewis 2000: 188). Sally's throwing a rock caused the window to break just in case there is a substantial range of not-too-distant alterations of her rock-throwing (including not throwing or throwing a heavier rock), which would have been followed by not-too-distant alterations of the window shattering (including not shattering or shattering into more pieces). My writing this paper did not cause the window to break since there is no such a range of alterations of my writing this paper which would have been followed by suitable alterations of the window breaking. This revision of counterfactual theory, generated partly in response to Schaffer's 'trumping preemption' cases, does *not* rely on the missing events feature or the delayed/changed effect feature of many cases of preemption mentioned earlier (not found in Schaffer's cases). Lewis resolves the trumping cases as follows: altering the trumping cause, while holding the trumped cause fixed, would be followed by a suitable alteration in the effect, but that is not true of the trumped cause. Had the trumped cause been altered and the trumping cause left unchanged, there would have been no alteration in the effect.

In order to test Influence Theory against Case A, we must make Case A deterministic since Lewis assumes that all causation is deterministic in constructing

Influence Theory. To that end, suppose that all transmissions (in Figure 5.1) are deterministic and that the collision at time *t* at *d*/*g* deterministically guarantees that *d* does not transmit, but does not prevent *d*'s firing (a particle/energy is present at *d* at *t* but in the next unit of time that particle/energy disappears). Assume also that had either sequence occurred without the other, *e*'s firing would have been deterministically guaranteed. Call this Deterministic Case A. Deterministic Case A will be consistent with Influence Theory only if there is a substantial range of different not-too-distant alterations of *g*'s firing that map on to a range of different alterations of *e*. However, further modifications to Case A will guarantee that the latter is false:

Assume that (1) by law no variation in the time of *g*'s firing would have made a difference to the timing of *e*'s firing. Had *g*'s firing been later, *d* would have fired and transmitted its particle/energy and *e* would have fired at exactly the time it did. If *g* had fired earlier, *g* would not have transmitted but *d* would have when it did and *e* would have fired when it did. Also assume that (2) by law variations in the manner of *g*'s firing would not have made a difference. For example, we can suppose that if *g* had possessed a greater quantity of energy, it would have lost the extra energy in moving to node *e*. If *g* had had less energy, it would not have transmitted at all, but would not have inhibited *d*'s firing and *d* would have transmitted.

Under these conditions, Influence Theory fails to determine that *g*'s firing is a cause of *e*'s firing. *g*'s firing cannot be altered in such a way that those alterations map on to alterations of *e*'s firing.[9]

M-set analysis

Ramachandran (1997, 1998) revises counterfactual theory by way of what he calls D-sets and M-sets, defined as follows:

> A non-empty set of events, S, is a dependence set (or D-set) for an event *y*, where *y* is not a member of S, just in case: (D) if none of the events in S had occurred, then *y* would not have occurred. Any D-set for *y*, S, is a minimal dependence set (an M-set) for *y* if and only if it is also true that: (M) no proper subset of S is a D-set for *y*.
>
> (Ramachandran 1997: 270)

The first version of M-set theory that Ramachandran proposes brings into play the 'missing intermediary events' feature of some preempted processes:

> For any actual events *c* and *e*, *c* causes *e* if and only if [A] *c* belongs to an M-set for *e* and [B] there are no M-sets for *e*, *M* and *N*, such that *M* contains *c* and *N* differs only in that it has one or more non-actual events in place of *c*.
>
> (Ramachandran 1997: 274)[10]

Clause [B] does the real work in cases of preemption. Although both preempting and preempted causes will be members of M-sets for the effect, the trick is to show that the preempted cause does *not* satisfy [B], but the preempting cause does. That will be true if there are missing events in the preempted line. Using the Figure 5.2 example, we see that the set {*a*'s firing, *b*'s firing} is an M-set for *e*'s firing. *b*'s firing, the preempting cause, does meet clause [B] since we cannot replace *b*'s firing with a non-actual event such as *d*'s firing in that set and still have an M-set. However, the preempted cause, *a*'s firing, fails to satisfy clause [B]. {*d*'s firing, *b*'s firing} is an M-set, *N*, for *e*'s firing where *d*'s firing is a non-actual event that replaces *a*'s firing if we make *M* {*a*'s firing, *b*'s firing}.

The key question in Case A is whether or not the preempted cause fails clause [B]. Since there are no missing *events* in the preempted process, *a*'s firing will not fail clause [B]. *a*'s firing and *b*'s firing are each members of an M-set, {*a*'s firing, *b*'s firing}, for *e*'s firing. However, there is no M-set, *M*, containing *a*'s firing, for example {*a*'s firing, *b*'s firing}, and a further M-set for *e*'s firing, *N*, for example {*b*'s firing, *d*'s firing}, that differs from the former only in that it has one or more non-actual events. After all, *d*'s firing is an actual event.

A second version of M-set theory is devised by Ramachandran in response to a case from Noordhof (1998a: 458) in which the preempted process includes no missing events, but the final effect of the preempted chain occurs after the candidate effect (the *a* node is connected to the *c* node which is connected to the *d* node, and the *b* node is connected to the *d* node): *a*'s firing at time 0 causes *c*'s firing at time 2 and *b*'s firing at time 1 causes *d*'s firing at time 3. *d*'s firing, which occurs at time 3, is caused by *b*'s firing, which preempts *a*'s firing from causing *d*'s firing.[11] What Ramachandran calls 'Analysis #2' brings into play the 'delayed effect' feature of the preempted process in this case: [12]

> For any actual events *c* and *e*, *c* causes *e* iff (2.1) *c* belongs to a temporal M-set for *e*, and (2.2) *c* belongs to an M-set for *e*, *M*, such that for any M-set for *e*, *N*, that differs only in that it contains one or more events in place of *c*, at least one of the events replacing *c* is actual and belongs to a temporal M-set for *e*.
>
> (Ramachandran 1998: 466)[13]

A set of possible events, *S*, is a temporal D-set for *e* if it is true that if none of the events in *S* were to occur, then *e* would not occur at *t*, the time of *e*'s actual occurrence. A temporal M-set for *e* is a temporal D-set with no proper subsets that are temporal D-sets for *e* (Ramachandran 1998: 466). In Noordhof's case, Ramachandran argues that *a*'s firing cannot be shown to satisfy 2.2 by reference to its membership in {*a*'s firing, *b*'s firing}, considered as our candidate for *M*. That set, says Ramachandran, does not satisfy 2.2 since there is a set, {*c*'s firing, *b*'s firing}, that differs from it only by including an actual event, *c*'s firing, in place of the firing of *a*, that does not belong to a temporal M-set for *d*'s firing (Ramachandran 1998: 466).[14] Unfortunately, the same reasoning does not apply to Case A. In Case A, if we substitute either of the events, *c*'s firing or *d*'s firing, for *a*'s firing in {*a*'s firing, *b*'s firing} we end up with M-sets for *e* that differ only from the original set in that each contains one actual event in place of the firing of *a* but that event belongs to a temporal M-set for *e*. These alternative sets differ from the original only in that each contains an actual event in place of *a*'s firing that does belong to a temporal M-set for *d*'s firing.

Noordhof

Noordhof's revision of counterfactual theory (1999) begins with the thought that a preempting causal process would display probabilistic dependence in the counterfactual circumstance of the absence of the preempted cause. However, that alone does not distinguish between preempting from preempted causes – the latter generally display such dependence in the absence of the former. Further clauses are introduced. Given the complicated nature of Noordhof's account, I start with an approximation. This first approximation emphasizes the 'missing intermediary events' feature of some preempted processes:[15]

> For any actual, distinct events e_1 and e_2, e_1 causes e_2 iff there is a (possibly empty) set of possible events Σ such that (I) e_2 is probabilistically Σ-dependent on e_1, and (II) every event upon which e_2 probabilistically Σ-depends is an actual event.

Probabilistic Σ-dependence is defined initially as follows:

> For any events e_1 and e_2, and any set of events Σ, e_2 probabilistically Σ-depends on e_1 if and only if (i) if e_1 were to occur without any of the events in Σ, then p(e_2) would be at least x, (ii) if neither e_1 nor any of the events in Σ were to occur, then p(e_2) would be at most y, and (iii) $x > y$.
>
> (Noordhof 1999: 104)

In cases of deterministic preemption, an effect will probabilistically Σ-depend on its preempting cause if Σ includes the preempted cause (assessing the probability of the effect just before it occurs). But if Σ includes the preempting cause, effects will also probabilistically Σ-depend on their preempted causes. Clause (II), the 'actual events' clause, is intended to distinguish preempting from preempted causes, at least in cases of deterministic early and late preemption.[16] If Figure 5.2 represents deterministic early preemption and *b*'s firing is in Σ, for instance, then although *e*'s firing is probabilistically Σ-dependent on the preempted cause, *a*'s firing, it is also probabilistically Σ-dependent on *d*'s firing, the missing intermediary event. The same kind of thing is not true of *b*'s firing if *a*'s firing is in Σ (Noordhof 1999: 102–4).

This strategy will not work in Case A (Figure 5.1). First, the effect, *e*'s firing, is Σ-probabilistically dependent on *a*'s firing, the preempted cause, if Σ includes *b*'s firing. If *a*'s firing is to be excluded it must be by way of clause (II). The difficulty is that when assessing *a*'s firing, with *b*'s firing in Σ, *e*'s firing is definitely *not* Σ-probabilistically dependent on any missing event since there are no missing events in the preempted *a*-process to play that role. It might be objected that there is a missing event, the missing transmission of a particle/energy from d to *e* and that *e*'s firing is probabilistically dependent on that non-actual event. In fact, this missing transmission is not a missing event. First, if it were an event and it were to occur in the absence of the preempting line, it would occur at a certain time, presumably between *d*'s firing and *e*'s firing, but there would be no such temporally intermediate event in the absence of the preempting line. Second, if it were an event and it were to occur, that event would have been caused by *d*'s firing and caused *e*'s firing, but there would have been no such event causally between *d*'s firing and *e*'s firing in the absence of the preempting line. Had the preempted line occurred on its own, *d*'s firing would have caused *e*'s firing directly and not by way of some further event.

For probabilistic late preemption, Noordhof recasts the notion of probabilistic Σ-dependence (bringing into play the delayed effect feature of late preemption):[17]

> e_2 probabilistically Σ-depends upon e_1 if and only if (1) if e_1 were to occur without any of the events in Σ, then for some time *t*, it would be the case that, just before *t*, p(e_2 at *t*) $\geq x$, (2) if neither e_1 nor any of the events in Σ were to occur, then for any time *t*, it would be the case that, just before *t*, p(e_2 at *t*) $\leq y$, and (3) $x > y$.
>
> (Noordhof 1999: 109–10)[18]

A further clause, Clause (III), is also added (1999: 113): e_2 occurs at one of the times for which p(e_2 at *t*) $\geq x > y$.[19] Noordhof argues that the preempting cause, in a case of probabilistic late preemption, either fails clause (I) or (III).[20] Reading Figure 5.3 as in Noordhof (1999: 99) as probabilistic late preemption and assessing the probability of the effect just before it would have happened, he reasons that if the effect had a background chance of occurring anyway, 'it will still occur prior to *d*'s firing' and then its probability of occurring later is 0 ('the very same event can't occur twice'), in which case *a*'s firing fails clause (I) (Noordhof 1999: 113). If *e*'s firing has no background chance of occurring, then p(*e* at t_0) is not raised by the presence/absence of *a*'s firing since had *e* fired by that route it would have fired later, and, thus, clause (III) is not satisfied (Noordhof 1999: 113–14).

This reasoning does not help in Case A, at least if we suppose that there is no background chance of *e*'s firing earlier than when it did. In that event, *a*'s firing does not fail clause (I), but neither does it fail clause (III). With respect to *a*'s firing in Case A, one of the times at which p(*e* fires at *t*) $x > y$ is t_0, the time at which *e* actually occurred. If *e* fired as a result of the *a*–*e* process, it would have occurred at the same time that it actually did occur.

A fourth and final clause is added to deal with 'anti-catalyst' and 'catalyst' cases in which the preempting line acts not to inhibit the preempting line, but to slow it down or speed it up where it is true that had the preempted process not been so influenced, it would have given rise to the effect at just the time it actually occurred (Noordhof 1999: 115–16).[21] Clause (IV) is meant to rule out the preempted cause as a genuine cause in such cases: (IV) e_2 probabilistically A-time-depends on e_1 (Noordhof 1999: 116). e_2 probabilistically A-time-depends upon e_1 just in case there is a (possibly empty) set of possible events A such that if e_1 were to occur without any of the events in A, then it would be the case that P(e_2 at t_0, the actual time of e_2) $\geq x$, whereas without e_1 or the A events, it would be the case that p(e_2 at t_0) $\leq y$, and $x > y$ (Noordhof 1999: 116). What can be a member of A is restricted to only those events 'whose absence leave untouched the relationship between candidate cause, e_1, and the p(e_2 at the time it actually occurred)' (Noordhof 1999: 117). Anti-catalytic events in the preempting line get ruled out from being members of A. With such events not being members of A, the preempted cause cannot satisfy (IV).

Clause (IV) does not help in Case A since *g*'s firing will be a member of A. *g*'s firing does not satisfy the conditions for exclusion from A set down by Noordhof. More specifically, the following condition is not met: if *g*'s firing occurred without satisfying (I) to (III) with respect to the firing of *e*, then the *a–d* chain would not raise the probability of *e*'s firing.[22] In fact, that is not true because *g*'s firing is only a *probabilistic* inhibitor of the transmission from *d* to *e*. The *a–e* process is highly reliable and the probability that *g*'s firing will inhibit the *d–e* transfer is 50%. Hence, we can safely assume if *g*'s firing occurred – without satisfying (I) to (III) – then the *a–d* chain would still raise the chance of the firing of *e* at time t_0.[23] *g*'s firing is not excluded from A. In that case, *e*'s firing is probabilistically A-time-dependent on *a*'s firing, the preempted cause.

McDermott

McDermott (1995, 2002) offers a sufficient-condition account of causation, but his account is a variant of the counterfactual theory since he characterizes the notion of a sufficient condition in counterfactual terms. A sufficient condition for an event *e* is a condition on what happens at a point such that given its satisfaction *e* would have occurred whatever had happened at other points (McDermott 2002: 96–7). 'A *minimal* sufficient condition for *E* is a sufficient condition in which no conjunct could be replaced by a weaker condition on what happens at that point without losing sufficiency' (McDermott 2002: 96–7). If *c* and *e* are distinct actual events, *c* is a *direct cause* of *e* if and only if the occurrence of *c* satisfies a conjunct in some satisfied minimal sufficient condition for *e* (McDermott 2002: 97). Indirect causation is a matter of more than chains of direct causation. *c* causes *e* if and only if (i) there is a chain of direct causation linking *c* to *e*, (ii) *c* is essential to the production of *e*, via that channel, (iii) *c* is essential in the production of an event meriting the description '*e*', via that channel (McDermott 2002: 100).[24] Finally, for McDermott, events are extremely fragile, and causation is not extensional.[25]

Since McDermott assumes determinism in constructing his theory, we should test it against Deterministic Case A. And since *g*'s firing is a direct cause of *e*'s firing, we can focus on his account of direct causation. His account gives the right result for the preempting cause. Even if g's firing had occurred by itself, *whatever else had happened*, it would have lead to e's firing at the same time and in the same manner. Had g's firing occurred and d's firing (had/had not) occurred, for example, e's firing would have occurred in Deterministic Case A. g's firing is both a sufficient condition for and a minimally sufficient condition for e's firing. McDermott's account, however, does not give the right result for the preempted cause. Had d's firing occurred by itself, *whatever else had happened*, it would have led to e's firing at the same time and in the same manner. If the relevant actual events are d's firing, … , g's firing and e's firing, then had d's firing occurred and g's firing (had/had not) occurred, e's firing would have occurred. Recall that d's firing is just the presence of a particle with its energy at the d node and e's firing is the presence of a particle with energy at the e node. Whether or not the d particle collides with the g particle, a particle/energy will be transferred to the e node if a particle is transferred to the d node. d's firing is also a minimally sufficient condition for e's firing.

Process theories and hybrid theories

Before considering transference theory (which is, in some loose sense, a 'process' theory) I will consider a hybrid theory that combines counterfactual theory and process theory: Schaffer's causes-as-probability-raisers-of-processes theory. My main interest is in Schaffer's account of a process as a law-governed sequence. I will suggest that this notion of a process does not help in preemption cases like Case A. Seeing why not will help to motivate a singularist reading of transference theory and point us towards a conception of causation's singularist component that will help us with persistence preemption.

Schaffer

Schaffer combines probabilistic counterfactual theory and process theory. A cause is a probability raiser of a process (a PROP) for its effect (Schaffer 2001a: 75–92). Probability raising is analysed in counterfactual terms. Actual event c is a probability raiser of a distinct actual event e just in case ch(e)-at-$t_c = p$, and $\neg c \; \Box\!\!\rightarrow$ ch(e)-at-$t_{\neg c} < p$ (Schaffer 2001a: 77). A process is a lawful sequence, a sequence of events that are pairwise subsumed under a fundamental dynamical law (Schaffer 2001b: 18, n. 11). c is *directly* process-linked to e if and only if linked by such a law. More generally, c is process-linked to e if and only if there is a chain of direct process links between them (Schaffer 2001a: 78).[26] When these definitions are plugged into the PROP account, Schaffer states the result as follows:

> *Analysis 1 Interpreted*: C is a PROP for E if and only if (i) there is an extended event 'E-line' containing actual distinct events <C', $D1$, D_2, … , Dn, E> in pairwise nomic subsumption relations, (ii) there is an actual event C at t_C'

which is distinct from $D_1, D_2, \ldots, D_n$ and E (C may or may not be distinct from C'), (iii) ch(E-line)-at-$t_C = p$, and (iv) $\neg C$ $\Box\!\!\rightarrow$ ch(E-line)-at-$t_{\neg C} < p$.

(Schaffer 2001a: 85)

Schaffer argues that each preempting cause is an essential part of a process that includes the effect (Schaffer 2001a: 87).[27] To illustrate, Schaffer considers an example in which Pam throws a rock at a window, while Bob, a more reliable thrower, holds off throwing. Had Pam not thrown her rock, Bob would have thrown his rock. Pam's throw is a preempting cause of the window shattering, although it lowers the probability of that event given Bob's greater reliability. Given that Pam's throw preempts Bob's throw in shattering the window, there is the sequence of events, which are pairwise lawfully related, of her throwing the brick, her brick flying through the air, hitting the window, and shattering it. 'Pam's throw is part of this process and an essential part, since without it if the window gets shattered at all it will be by a different process entirely: Bob's process' (Schaffer 2001a: 87).

Again consider Case A. Assume that the processes involved are governed by fundamental probabilistic laws such that the events on the main line and on the alternate line are pairwise subsumable under fundamental probabilistic laws. By law, *d*'s firing increases the probability of *e*'s firing and, by law, *g*'s firing increases the chances of *e*'s firing. Schaffer's lawful-sequence account of processes gives, then, the wrong result for the preempted cause. *d*'s firing is an essential part of a Schaffer-process culminating in *e*'s firing. This becomes even clearer when we consider Schaffer's solution to a case in which a major's order to a corporal trumps a sergeant's equivalent order. 'Applying the lawful sequence analysis of processes: there are trumping laws linking ranking orders to decisions. Since the major's order is the ranking order, only the major's order, not the sergeant's, instantiates the antecedent. The corporal's decision instantiates the consequent' (Schaffer 2001a: 87). The sergeant's order is not, but the major's order is, on process to the effect because the former does not instantiate an antecedent of such a (fundamental) law but the latter does.[28] The same approach does not work for the preempted cause, *d*'s firing (Figure 5.1), since it meets the conditions of 'Analysis 1 Interpreted':

(i) there is an extended event 'E-line' containing actual distinct events <*d*'s firing, *e*'s firing> in pairwise (probabilistic) nomic subsumption relations, (ii) there is an actual event *d*'s firing at t_d's firing which is distinct from *e*'s firing, (iii) ch(E-line)-at-$t_c = p$, and (iv) $\neg d$'s firing $\Box\!\!\rightarrow$ ch (E-line)-at-$t_{\neg c} < p$.[29]

d's firing is an essential part of a process that includes *e*'s firing under the nomic conception of processes.

The lesson is that the lawful-sequence conception of a 'process' is not the way to go to handle persistence preemption. I will now consider a simple form of transference theory with that lesson in mind. As we shall see, for transference theory to handle Case A, the transference relation must not be just a matter of lawful sequence.

Transference theory

Consider a simple, very broad form of transference theory, not attributable to any proponent of transference theory, beginning with the broad part. A narrow form of transference theory requires cross-object transfers for causation, ruling out causation within a single object (Aronson 1971). To avoid this limitation, broaden the notion of transfer to include transfers across spatial locations or even 'transfers' from one temporal part of an object to another temporal part of that same object (Dowe 2000b: 54). Also assume that the theory is simple with the only clause being the following: *c* causes *e* just in case there is a transfer of energy/momentum from the *c*-object/location/temporal part to the *e*-object/location/temporal part, with causes and effects consisting of the manifestations of energy/momentum. Finally, assume that the quantity 'transferred' persists, at least in part, through the transfer. There is identity (or perhaps, partial identity) over time of the quantity that is transferred. A preempting cause is, then, distinguishable from a preempted cause in virtue of the fact that the energy/momentum of the effect event is traceable to the preempting cause event, but not to the preempted cause event. In Case A, the causal facts can be read off from the energy transfers across locations/temporal parts.

The advantage that this form of transference theory has over counterfactual theory in Case A is that the former posits something, a quantity of energy/momentum, that persists from cause to effect, but counterfactual theory does not. However, for this advantage not to be illusory – for transference theory really to render the correct verdict in Case A – the persistence of the relevant quantity of energy cannot just be a matter of spatiotemporal or nomological relations among temporal parts of that quantity.[30] Otherwise, the difficulty of distinguishing preempting from preempted returns. If the persistence of a quantity of energy is just a matter of the temporal stages (of that quantity) standing to each other in certain spatiotemporal relations, transference theory cannot distinguish between *d*'s firing and *g*'s firing, standing as they do in the same temporal/distance relations to *e*'s firing. Each quantity would have an equal claim to persistence. Each firing would have an equal claim to being a cause of *e*'s firing. Similarly, if the persistence of *q* is just a matter of nomological relations between successive temporal stages the same trouble reemerges (as it does for Schaffer processes). As we saw, *d*'s firing and *g*'s firing stand in the same law-governed probabilistic relations to *e*'s firing. The probabilistic laws do not distinguish between them. And if the persistence of this quantity is a matter of causal relations among temporal parts then the theory will be circular. For transference theory to give the right verdict, the persistence of this quantity must be a matter of that quantity (or a portion thereof) being wholly present at each moment that it exists. I am not suggesting that is how transference theorists in fact think about the persistence of energy/momentum, but only that that is required if transference theory is to deal with Case A.

In this simple form, transference theory is, thus, more successful with Case A than is counterfactual theory. The key difference is that transference theory, as interpreted here, tells a certain kind of story about the singularist processes which connect causes with their direct effects, positing a non-causal physical mechanism for carrying direct causal influence. Under this interpretation, transference theory is just one example of a theory, not the one I support, that posits such a singularist component to causation (and, as it happens, no generalist component). Such a theory need not be singularist as a whole, but will include such a component. Transference theory, as interpreted here, provides a useful example of such a theory to contrast with counterfactual theory. How should this non-causal mechanism be characterized more generally? There is a local mechanism for the transmission of direct causal influence which involves a persisting 'entity' of some sort where the *persistence* of this 'entity' cannot be analysed as a matter of spatio-temporal or nomological relations among temporal parts of that entity since either of these options will resurrect the difficulty of distinguishing the preempting from the preempted cause in Case A. In addition, since an analysis of causation must not make use of concepts which themselves are analysed causally, the entity's persistence must not be a matter of a series of causally connected temporal stages. More generally, the persistence of this 'entity' is not a matter of temporal parts at all, if it is to help with cases like Case A – this 'entity' (or a portion thereof) is wholly present at each moment of its existence (a three-dimensionalist view).

I have suggested an alternative account of causation's singularist/persistence component as the persistence of properties – understood as tropes (Ehring 1997). More specifically, I argued that causal relata are tropes and that, roughly, causes and effects are connected by persisting tropes. In more detail, I first defined a notion of being strongly causally connected as follows. Tropes P and Q are strongly causally connected if and only if:

(1) P and Q are lawfully connected, and either
(2) P is identical to Q or some part of Q, or Q is identical to P or some part of P, or
(3) P and Q supervene on tropes P' and Q' which satisfy (1) and (2).

Clause (1) is a placeholder for causation's generalist component. Clause (3) is a placeholder for some relation meant to handle nonreducible properties, which I now think should be worked out by way of the part-whole relation as applied to types, understood as classes of tropes (Ehring forthcoming). Clause (2) is causation's singularist component. Causes are connected to their direct effects by way of persisting or partially persisting tropes. The most basic form that this persistence takes is that of an individual trope persisting unchanged. Other forms include partial destruction of a trope, trope fission and trope fusion. Since strong causal connection is a symmetrical relation, it must be supplemented with an account of causal priority to guarantee that causation is an asymmetrical relation (see Ehring 1997 for an account of causal asymmetry). Trope P at t causes trope Q at t' if (A) P at t is strongly causally connected to Q at t', and P at t is causally prior to Q at t'. Since events connected by a chain of indirect causation may fail to satisfy this condition, a second sufficient condition is required: (B) there is a set of properties ($R_1, \ldots, R_n$) such that P is a cause of R_1 under clause (A), … , and R_n is a cause of Q under clause (A). P at t causes Q at t' if and only if either (A) or (B) is true. In Case A, the relevant trope is the quantitative energy property/trope of the particle that moves along the *b*-route. That trope persists from the *b* node to the *e* node. There is no comparable persistence along the *a*-route through to *e*.

Summary

I have argued that a theory that includes a singularist component of causation involving persistence of some 'entity' does a better job with Case A than a number of counterfactual theories. Demonstrating this comparative advantage with respect to Case A, however, is not a full defence of the former type of theory nor of the version I prefer. A full defence requires an argument for the claim that causal relata are tropes, an argument for the claim that some tropes can persist and a discussion of apparent cases of causation in which there is no persistence, including apparent cases of causation by and of omissions (see Ehring 1997 for some of that defence).

Notes

1 A singularist connection between cause and effect is a process or relation the realization of which does not depend upon what happens in other regions or on how other token events of the same type are related.
2 After a node acquires a particle/energy at a time *t*, if the transmission does not take place in the next unit of time, then the particle/energy disappears completely in that next unit of time.
3 Schaffer agrees that there is no intrinsic difference, but argues that there is an extrinsic difference (which spell came earlier) that in combination with the law that (i) the first spell comes true determines a causal difference. McDermott, however, rejects this appeal to (i) to support the preemption verdict. McDermott argues that, given the Lewis–Ramsey model of laws, if (i) is a law then so is the logically equivalent proposition that (ii) when nonequivalent spells are cast the first comes true, but when only equivalent spells are cast the last comes true. Using (ii) as a guide would make Morgana efficacious. McDermott also argues that Schaffer's appeal to considerations of simplicity will not have the consequence that (i) is a law, but (ii) is not, since both are theorems and, on that model, simplicity considerations apply to axioms, not theorems.
4 Or more precisely, this example is a deterministic version of a indeterministic case from Menzies (1989: 646) which itself is an indeterministic modification of a case from Lewis (1986: 200). Figure 6 is borrowed from Menzies (1989: 646).
5 Had the preempted line not been blocked, it would have given rise to the effect at the time the actual effect occurred by way of some additional events (*missing intermediary events*).
6 Menzies (1989: 645–7) discusses a probabilistic version of early preemption and shows that a probabilistic version of Lewis's original formulation is inadequate in that there can be a chain of probabilistic dependence, as the latter is characterized by Lewis, without causation. Menzies offered a revision of counterfactual theory which requires that causal chains be spatiotemporally continuous in some sense to deal with this problem. Menzies later (1996) gave this account up because it rules out temporal action at a distance and because it cannot handle late preemption.

7 I borrow Figure 5.3 from Menzies (1989: 652).

8 Ganeri, Noordhof and Ramachandran (1996: 223) complain that a ‘quasi-dependence’ approach rules out brute singular causation in cases of preemption. Noordhof (1999: 101) also complains that this ‘quasi-dependence’ approach misclassifies the preempted cause in certain cases of probabilistic preemption.

9 This variation in the case is based on cases found in Collins (2000: 231), Schaffer (2001b: 16) and McDermott (2002: 92).

10 ‘For any pre-empted cause, *x*, of an event, *y*, there will be at least one possible event … which fails to occur in the actual circumstances *but which would have to occur in order for x to be a genuine cause of y* … All genuine causes, on the other hand, *do* seem to run their full course; indeed, they presumably count as genuine precisely because they do so’ (Ramachandran 1997: 273).

11 It takes two units of time to travel from node to node.

12 Noordhof (1998a) also suggests that M-set analysis will fail for indeterministic causes of effects that have a probability of occurring anyway.

13 Ramachandran (1998) proposes more than one revised M-set account.

14 Cases of frustration involve no ‘missing events’, but they do involve ‘delayed effects’, and if ‘Analysis #2’ gets a grip, it does so, in part, because of this feature. Ramachandran (1998: 467) does not think this solution will work in frustration cases in which the connection between *a*’s firing and *d*’s firing is direct.

15 Noordhof’s account descends from an account offered by Ganeri, Noordhof and Ramachandran (1996, 1998).

16 ‘All it relies upon is the existence of possible events which were actually suppressed due to preemption’ (Noordhof 1999: 104).

17 With probabilistic late preemption the preempted cause while raising the probability of the effect does not raise it *at the time the effect occurred* but only later.

18 This modification also incorporates Noordhof’s conviction that the presence of a cause makes the probability of the effect greater at the time of its occurrence than at any other time if the cause were not present relative to the events in Σ.

19 Clause (II) is also modified in response to cases of probabilistic early preemption in which the alternate line is blocked, not by the main line, but by a distinct inhibitory causal process and in which the preempted cause cannot be ruled out by clause (II). The modified clause reads as follows: (II)' for any superset of Σ, Σ*, (where $\Sigma \subseteq \Sigma^*$), if e_2 probabilistically Σ*-depends upon e_1, then every event upon which e_2 probabilistically Σ*-depends is an actual event (Noordhof 1999: 107).

20 On the other hand, Noordhof argues that the firing of *b* satisfies clauses (I–III) (1999: 113).

21 The four clauses mentioned form a necessary condition for causation to be supplemented with an account of causal asymmetry (Noordhof 1999: 120).

22 *g*’s firing is excluded from *A*: (a) If *g*’s firing is a member of A, then <*a*’s firing, *e*’s firing> satisfies (IV); and (b) If *g*’s firing is not a member of A and we replace (IV)(1) and (2) with (1*) If *a*’s firing and *g*’s firing were to occur, with none of the events in A occurring, nor *g*’s firing satisfying any of (I) to (III) regarding *e*’s firing, then it would be the case that p(e_2 at t_0) $\geq x$; (2*) If *g*’s firing were to occur with neither *a*’s firing nor any of the events in A occurring, nor *g*’s firing satisfying any of (I) to (III) regarding *e*’s firing, then it would be the case that p(e_2 at t_0) $\leq y$ then <*a*’s firing, *e*’s firing > would not satisfy (IV) (Noordhof 1999: 117). In his (2000: 323), Noordhof weakens this test: ‘I need not have formulated the test condition … in terms of an event failing to satisfy *any* of conditions (I) to (III). Instead, I might have required just that the event fail to satisfy *all* of condidions (I) to (III).’

23 If that is not true, we can tweak the case by increasing the reliability of the *a*-process and lowering the chance of *g*’s firing inhibiting the *a*-process.

24 Clause (ii) is spelled out as follows: without *c*, *e* might not have occurred, or without *c*, *e* would have occurred via another channel (its occurrence would have been dependent on some satisfied condition *X* which 'contributed nothing' to the actual chain of direct causation linking *c* to *e*) (McDermott 2002: 99).

25 Clause (iii) is designed to narrow down the analysis to cover the relation of causing and not the relation of causing-or-affecting, which is defined by clauses (i) and (ii) (McDermott 2002: 100).

26 When a thrown brick breaks a window there is a process consisting of 'a sequence of events from the throwing of the brick, through its intermediary trajectories, to the shattering of the window; and the events in this sequence will be pairwise lawfully related (and not merely covariants, and rightly temporized)' (Schaffer 2001a: 78).

27 'If the *E*-process includes *C*, then *C* is a PROP for *E* if and only if *C* is an essential part of the *E*-process: without *C*, if E still occurs it is via a different process entirely, rather than the same process slightly altered' (Schaffer 2001a: 87).

28 Schaffer does not assume that all processes are governed by nonprobabilistic laws. See his probabilistic version of the sergeant case (Schaffer 2001a: 80).

29 Schaffer also offers an alternative, but related, analysis such that *c* is a PROP for *e* if and only if *c* is a continuous PROP for *e*. For details see Schaffer 2001a: 90. But as Schaffer points out there is a significant disadvantage to this account. It rules out the possibility of spatiotemporally discontinuous causation.

30 At a minimum, the *denial* that energy or momentum can have identity over time seems to mean that transference theory, even under our very broad reading, would have no chance of getting the causal facts right in Case A.

Chapter 6

Probability and causation

Michael Tooley

What conceptual connections, if any, are there between causation and probability? This question raises many issues, and I shall certainly not attempt to address all of them here. In particular, a full account of the relations between probability and causation would need to cover both causal laws and causal relations between states of affairs, and in the present discussion I shall ignore the former and confine myself to the question of what conceptual connections there are between probability, variously conceived, and the relation of causation.

Four answers to the latter question are, I think, especially important, and it is upon these that I shall focus. The first two answers share the view that the concept of the relation of causation can be analysed probabilistically, but they disagree with regard to what the relevant concept of probability is. According to the first, causal relations between states of affairs can be reduced to non-causal facts, including facts about *relative frequencies*; according to the second, causal relations are reducible, instead, to states of affairs that involve *objective chances*.

The third main view, by contrast, holds that causal relations are not reducible to non-causal facts. But it also holds that the analysis of the concept of causation does involve the idea of probability – specifically, that of logical probability. So there is a conceptual connection between causation and probability, but not of a sort that generates a reductionist account of causation.

Finally, there is the view that – while there can, of course, be probabilistic causal laws – it is not the case, contrary to the first three views, that the relation of causation itself is to be analysed in any way that involves any concept of probability.

Of these four views, the third is, I believe, the correct one. But any attempt to establish that it is so would require a very extended discussion, since one needs to argue both that the concept of the relation of causation cannot be analytically basic, and that no analysis of the concept of causation that does not involve the concept of logical probability can possibly be sound. My goal here, accordingly, will be the less ambitious one, first, of arguing that reductionist accounts of causation in terms of either relative frequencies, or objective chances, together with other non-causal facts, cannot be sound, and second, of making plausible the claim that there are necessary connections between causation and logical probability.

1 Reductionist versus realist approaches to the relation of causation

Two distinctions will be important for the discussion. The first is that between realist and reductionist approaches to causation. So let us briefly consider that distinction.

Reductionists with regard to the relation of causation claim that causal relations between states of affairs are reducible to non-causal facts, including non-causal properties of, and relations between, states of affairs, whereas realists claim that no such reduction is possible. But what is the relevant concept of reduction here? The answer is that reduction can take two forms. On the one hand, there are analytical reductions, where the relations in question hold as a matter of logical necessity, broadly understood. On the other, there are reductions that involve a contingent identification of the relation of causation with some complex relation that is analysable completely in non-causal terms.

In this essay, however, the question is whether there is any necessary relation between probability and causation, so we can confine our attention to the idea of an analytic reduction. Let us consider, then, how analytic reductionist accounts of causation are best characterized. A traditional way of formulating things is in terms of whether the concept of causation is *analysable* in non-causal terms. It seems preferable, however, to employ, instead, the slightly broader concept of *logical supervenience* – a concept that can be explained as follows.

Let us say that two worlds, W and W^*, agree with respect to all of the properties and relations in some set, S, if and only if there is some one-to-one mapping, f, between the individuals in the two worlds, such that (1) for any individual x in world W, and any property P in set S, x has property P if and only if the corresponding individual, x^*, in W^*, also has property P, and vice versa, and (2) for any n-tuple of individuals, $x_1, x_2, \ldots x_n$ in W, and any relation R in set S, $x_1, x_2, \ldots x_n$ stand in relation R if and only if the corresponding individuals, $x_1^*, x_2^*, \ldots x_n^*$, in W^*, also stand in relation R, and vice versa. Then we can say that the properties and relations in some other set T that is completely distinct from S are logically supervenient upon the properties and relations in set S if and only if, for any two worlds W and W^*, if W and W^* agree with respect to the properties and relations in set S, they must also agree with respect to the properties and relations in set T.

Given the concept of supervenience, two slightly different, general forms of reductionism with respect to causal relations can be set out, of which the stronger, and perhaps more common form, is this:

> *Strong Reductionism with respect to Causal Relations*
> Any two worlds that agree with respect to all of the non-causal properties of, and relations between, particulars, must also agree with respect to all of the causal relations between states of affairs.

According to this first version of reductionism, then, causal relations are logically supervenient upon non-causal properties and relations of particulars. But this strong form of reductionism with respect to causal relations may be exposed to objections that one might well prefer to avoid, since, unless one is prepared to defend a singularist conception of causation, according to which causal relations between states of affairs need not fall under any laws, the above strong reductionist thesis can only be true if the following reductionist thesis is also true:

> *Reductionism with respect to Laws*
> Any two worlds that agree with respect to all of the non-causal properties of, and relations between, particulars, must also agree with respect to all laws.

This latter thesis, however, is very problematic. In particular, there are strong arguments for the view that it is logically possible for there to be basic laws that are uninstantiated, and therefore that it is logically possible for there to be two worlds that, although they agree with respect to all of the non-causal properties of, and relations between, particulars, do not agree with respect to all laws, since they disagree with respect to at least one basic law that is present in the one world – even though it has no instances – but not present in the other.

One argument, for example, in support of the possibility of basic, uninstantiated laws may be put as follows. Suppose, for the sake of illustration, that our world contains psychophysical laws according to which various types of brain states causally give rise to emergent properties of experiences. Let us suppose, further, that at least some of these psychophysical laws connecting neurophysiological states to phenomenological states are basic – that is, incapable of being derived from any other laws, psychophysical or otherwise – and, for concreteness, let us suppose that the psychophysical law connecting a certain type of brain state to experiences involving a specific shade of purple is such a law. Finally, let us assume that the only instances of that particular law at any time in the history of the universe involve sentient beings on Earth. Given these assumptions, consider what would have been the case if our world had been different in certain respects. Suppose, for example, that the Earth had been destroyed by an explosion of the sun just before the point when, for the first time in history, a certain sentient being would have observed a purple flower, and would have had an experience with the corresponding emergent property. What counterfactuals are true in the alternative possible world just described? In particular, what would have been the case if the sun had not gone supernova when it did? Would it not then have been true that the sentient being in question would have looked at a purple flower, and thus have been stimulated in such a way as to have gone into a certain neurophysiological state, and then to have had an experience with the relevant emergent property?

It seems to me very plausible that the counterfactual in question is true in that possible world. But that counterfactual cannot be true unless the appropriate psychophysical law obtains in that world. In the world where the sun explodes before any sentient being has looked at a purple flower, however, the law in question will not have any instances. So if the counterfactual is true in that world, it follows that there can be basic causal laws that lack all instances.

It is sometimes suggested that the problem posed by basic laws that lack instances can be accommodated in a satisfactory way if one has an ontology in which dispositions, propensities, and objective chances are themselves ultimate properties, rather than being reducible to laws plus non-dispositional properties. But this is not a satisfactory solution, for at least two reasons. In the first place, there are, as we shall see, very serious objections to the idea of ontologically ultimate dispositions, propensities and objective chances. In the second place, while it may be possible to handle some uninstantiated basic laws by appealing to ontologically ultimate dispositions, propensities and objective chances, it seems clear that this cannot always be done. For suppose that it is an uninstantiated basic law that something's having property *A* gives rise, at the appropriate temporal distance, to an instance of some other property, *B*. If *A* is, say, a conjunction of two properties, *C* and *D*, it may be that both of those properties are instantiated – in different locations – and that one can say that instances of *C* have the dispositional property of giving rise to instances of *B*, if *D* is present. But if *A* is, instead, a *simple* property, then, given that there are no instances of *A*, there is nothing that can have any relevant dispositions or objective chances to give rise, at the appropriate temporal distance, to an instance of *B*.

Given, then, that the thesis of Reductionism with respect to Laws appears untenable, if one is not prepared to embrace a singularist conception of causation, and to hold that causal relations between states of affairs can exist in the complete absence of any relevant laws, one will probably want to shift from the thesis of Strong Reductionism with respect to Causal Relations to the following, more modest thesis:

> *Moderate Reductionism with respect to Causal Relations*
> Any two worlds that agree both with respect to all of the non-causal properties of, and relations between, particulars, and with respect to the truth of all law statements that involve no causal concepts, must also agree with respect to all causal relations between states of affairs.

For this thesis, by allowing one to combine reductionism with regard to causation with a realist view of laws – such as, for example, the view that laws are certain second-order relations between universals – enables one to avoid having to attempt to answer the very strong objections that have been directed against reductionist accounts of laws of nature.[1]

Given this distinction between strong and moderate reductionism with regard to causation, I can now offer a more precise formulation of the views that I shall be arguing for here: my goal is to show that the attempt to offer a reductionist account of causation in terms of non-causal states of affairs, including ones that involve either relative frequencies or objective chances, fails not only if one is putting forward the strong reductionist thesis, but also if one is opting instead for the moderate reductionist programme. So the probabilistic reductions fail, and they do so even if one combines a reductionist account of causation with a realist view of laws.

2 Humean versus non-Humean reductionism

The other distinction that is important for the present discussion is that between what may be called Humean and non-Humean states of affairs. So let us consider that distinction.

A principle that Hume frequently appealed to was that there could not be logical connections between distinct existences. But what exactly are distinct existences? An existence, here, I think, is best viewed as a state of affairs. If so, the question is when two states of affairs are distinct. One answer that might be offered is that two monadic states of affairs, such as *a*'s having property *P* and *b*'s having property *Q*, are distinct if and only if either *a* is not identical with *b* or *P* is not identical with *Q*. But that analysis does not seem right, since distinct existences, so understood, could very well be logically related. In particular, if *b* were a part of *a*, then *b*'s having a certain property might very well entail *a*'s having a certain property.

The natural reaction to this problem is to shift from talk about things' not being identical to talk about things' not overlapping: two monadic states of affairs, such as *a*'s having property *P* and *b*'s having property *Q*, are distinct if and only if either *a* and *b* do not overlap, or else properties *P* and *Q* do not overlap. (Overlap of properties would need, of course, further explanation, but the basic thought here is that if there are conjunctive properties, then any such property overlaps each of its conjuncts.)

The idea now is to explain the distinction between Humean and non-Humean states of affairs along roughly the following lines. First, any property or relation with which one can be directly acquainted – that is 'directly observable', which is 'immediately given' in experience – is *ipso facto* a Humean property or relation. Second, any state of affairs all of whose constituent properties and relations are Humean is a Humean state of affairs. Finally, any other state of affairs, *S*, is Humean if and only if there is no set, *C*, of Humean states of affairs such that *C* together with *S* entails the existence of a state of affairs, *T*, that is distinct from *S*, and that is not entailed by *C* alone.

Finally, given the concept of a Humean state of affairs, a reduction may be classified as Humean if the relevant reduction base consists entirely of Humean states of affairs. Otherwise, the reduction is non-Humean.

Why is this distinction important? The answer is connected with the choice between reductionist approaches to causation and those realist approaches that do not view causation as directly observable, and with the reasons that philosophers have often found approaches of the latter sort problematic. Historically, the objections to such approaches were twofold. First, there was the semantical problem of even making sense of such accounts of causation, given that philosophers had no viable account of the meaning of theoretical terms, realistically interpreted.

Second, there was the epistemological problem of how one could justify beliefs about states of affairs that, by definition, were not reducible to ones that were open to observation.

The semantical problem was completely disposed of with the development, by philosophers such as R. M. Martin (1966) and David Lewis (1970), of satisfactory accounts of the meaning of theoretical terms, realistically interpreted, while, as regards the epistemological problem, most philosophers have come to accept the idea that something along the lines of an abduction/hypothetico-deductive method/inference to the best explanation can provide a satisfactory account of the justification of beliefs about theoretical entities. The details, however, are by no means settled, and some philosophers – most notably, Bas van Fraassen (1989) – believe that the whole idea of inference to the best explanation is unsound.

Given these two developments, why have theoretical-term, realist approaches to causation remained relatively unexplored? The answer, at least in a large part, I think, is that such approaches typically involve the postulation of non-Humean states of affairs – such as second-order states of affairs, involving relations between universals – that logically entail the existence of corresponding regularities.

Because of this, it is very important to distinguish between reductionist accounts of causation that are Humean and those that are not, for if the crucial objection to theoretical-term, realist approaches to causation is that there cannot be logical connections between distinct states of affairs, then non-Humean reductionist approaches are problematic in precisely the same way. In one very fundamental respect, therefore, non-Humean reductionist approaches to causation are much closer to realist approaches than they are to Humean reductionist accounts, and this is philosophically very important.

It is crucial to ask, therefore, of any particular reductionist account of causation, whether it is Humean or non-Humean. In the case of a reductionist account of causation that is formulated essentially in terms of relative frequencies, and that either makes no use of the idea of laws, or else uses the concept of laws of nature but analyses that concept in reductionist fashion, one has a Humean account. But if the account involves relative frequencies, together with a realist view of laws the account is non-Humean. Similarly, reductionist accounts that make use of objective chances may be Humean or non-Humean, depending upon what account is offered of objective chances. If the latter are explained in terms of non-causal laws of nature, together with categorical properties plus relations, and if a reductionist view of laws is adopted, then one has a Humean account. But if, instead, a realist view of laws of nature is employed, or if, as is generally the case, objective chances are treated as ontologically ultimate, then one has a non-Humean, reductionist account of causation.

3 Causality and relative frequencies

Through the end of the nineteenth century, almost all philosophers thought of causation as being connected with conditions that were totally sufficient to ensure the occurrence of an event. But that changed in the twentieth century, with the emergence of quantum physics, and the development of the social sciences, and many philosophers gradually came to feel that causation is not restricted to cases where there are causally sufficient conditions for the occurrences of events.

What implications does this have for the philosophy of causation? One possible view is that it has very little relevance. For while quantum physics certainly appears to provide excellent reason for holding that causation can be present in situations that do not fall under deterministic laws, this need not imply that one's *concept* of causation has to be revised. Perhaps all that is needed is a concept of probabilistic laws, a concept which can then be combined with one's prior concept of causation to generate a satisfactory account of causation in probabilistic settings.

But there is also a very different possibility that needs to be explored: perhaps the right route involves an account of causation that is itself genuinely probabilistic, so that the concept of probability, rather than merely entering via probabilistic laws, is part of the very analysis of the relation of causation itself.

3.1 The basic approach

The earliest attempts to formulate a probabilistic analysis of causation were advanced by Hans Reichenbach (1956), I. J. Good (1961, 1962) and Patrick Suppes (1970), and they all were based upon the idea of probability understood in terms of relative frequency. Moreover, no use was made of the idea of laws of nature, let alone of a realist conception of laws, so that what Reichenbach, Good and Suppes offered were strong reductionist accounts of causation, and ones that involved only Humean states of affairs.

At the heart of any probabilistic analysis of causation is the idea that causes must, in some way, make their effects more likely, and within these initial probabilistic accounts of causation the basic idea was to analyse what it is for a cause to make its effect more likely in terms of the notion of *positive statistical relevance*, where an event of type B is positively relevant to an event of type A if and only if the conditional probability of an event of type A relative to an event of type B is greater than the unconditional probability of an event of type A. Thus Suppes (1984: 151), for example, introduces the notion of a prima facie cause, defined as follows: 'An event B is a prima facie cause of an event A if and only if (i) B occurs earlier than A, and (ii) the conditional probability of A occurring when B occurs is greater than the unconditional probability of A occurring.'

Perhaps the most crucial test for any theory of causation is whether it can provide a satisfactory account of the direction of causation. What account can be offered, given a probabilistic approach? One possibility, of course, is to incorporate the 'earlier than' relation into one's analysis of causation, and to use that relation to define the direction of causation – as was done, for example, by Suppes (1970).

It is widely thought, however, that this is not satisfactory. One reason is that it then follows immediately both that it is logically impossible for a cause and its effect to be simultaneous, and for a cause to be later than its effect, and while both things may be the case, the fact that many people have thought, for example, that time travel into the past is logically possible surely provides good reason for holding that it cannot be an *immediate* consequence of the analysis of causation that backward causation is logically impossible.

Another consideration is that there is a serious problem about what it is that forms the basis of the direction of time, and a causal theory of time has been thought by many philosophers to be a possibility worthy of serious consideration. If so, then the direction of causation cannot be defined in terms of the direction of time.

Because of considerations such as these, most advocates of a probabilistic approach to causation have wanted to analyse the direction of causation in probabilistic terms. What are the prospects for doing this? The first thing to note is that the postulate that a cause raises the probability of its effect does not itself provide any direction for causal processes. For when the following equation for conditional probabilities:

$$\text{Prob}(E/C) \times \text{Prob}(C) = \text{Prob}(E \;\&\; C) = \text{Prob}(C/E) \times \text{Prob}(E)$$

is rewritten as:

$$\text{Prob}(E/C)/\text{Prob}(E) = \text{Prob}(C/E)/\text{Prob}(C)$$

one can see that $\text{Prob}(E/C) > \text{Prob}(E)$ if and only if $\text{Prob}(C/E) > \text{Prob}(C)$. So causes raise the probabilities of their effects only if effects also raise the probabilities of their causes.

How, then, can the direction of causation be analysed probabilistically? The most promising suggestion was set out by Reichenbach in his book *The Direction of Time* (1956). Reichenbach's proposal involves the following elements: first, what he referred to as 'the Principle of the Common Cause'; second, a probabilistic characterization of a 'conjunctive fork'; third, a proof that correlations between event-types can be explained via conjunctive forks; and, fourth, a distinction between open forks and closed forks.

As regards the first element, Reichenbach's Principle of the Common Cause is as follows: '*If an improbable coincidence has occurred, there must exist a common cause*' (Reichenbach 1956: 157). Here the basic claim is that if events of type *A*, say, are more likely to occur given events of type *B* than in the absence of events of type *B*, and if the explanation of this is not that events of type *A* are caused by events of type *B*, or vice versa, then there must be some third type of event – say, *C* – such that events of type *C* cause both events of type *A* and events of type *B*.

Second, there is Reichenbach's characterization of the idea of a conjunctive fork, which – using a slightly different notation – can be set out as follows (1956: 159):

Events of types *A*, *B*, and *C* form a conjunctive fork if and only if:

(1) Prob(A & B/C) = Prob(A/C) × Prob(B/C)

(2) Prob(A & B/not-C) = Prob(A/not-C) × Prob(B/not-C)

(3) Prob(A/C) > Prob(A/not-C)

(4) Prob(B/C) > Prob(B/not-C)

Third, Reichenbach then shows that, provided that none of the relevant probabilities is equal to zero, equations (1) through (4) entail:

(5) Prob(A & B) > Prob(A) × Prob(B)

This in turn entails:

(6) Prob(A/B) > Prob(A)

(7) Prob(B/A) > Prob(B)

So we see that the existence of a conjunctive fork involving event types A, B and C provides an explanation of a statistical correlation between the event-types A and B.

Finally, Reichenbach then distinguishes between open forks and closed forks. Suppose that events of types A, B, and C form a conjunctive fork, and that there is no other type of event – call it E – such that events of types A, B and E also form a conjunctive fork. Then A, B and C form an open fork. On the other hand, if there is another type of event, E, such that events of types A, B and E also form a conjunctive fork, what one has is a closed fork.

As Reichenbach emphasizes, there can certainly be conjunctive forks that involve common effects, rather than common causes (1956: 161–2). But since conjunctive forks can, as we have just seen, explain statistical correlations, if there were an *open* fork that involved a common effect, then the relevant statistical correlation would be explained, even though there was no common cause, and this would violate the Principle of the Common Cause. Hence, conjunctive forks involving a common effect must, if Reichenbach is right, always be closed forks. All open forks, therefore, must involve a common cause, and so the direction of causation is fixed by the direction given by open forks.

3.2 Objections

This is a subtle and ingenious attempt to offer a probabilistic analysis of the relation of causation, and one that appeals only to Humean states of affairs. Unfortunately, it appears to be open to a number of decisive objections.

3.2.1 Accidental, open forks involving common effects

The basic idea here is simply this. Suppose that *A* and *B* are types of events that do not cause one another, and for which there is no common cause. Then it *might* be the case that the conditional probability of events of type *A* given events of type *B* was *exactly equal* to the unconditional probability of events of type *A*, but surely this is not necessary. Indeed, it would be more likely that the two probabilities were at least slightly different, so that the conditional probability of events of type *A* given events of type *B* was either greater than or less than the unconditional probability of events of type *A*.

Let us suppose, then, that the second of these alternatives is the case. Suppose, further, that the occurrence of an event of type *A* is a causally necessary condition for the occurrence of a slightly later event of type *E* and, similarly, that the occurrence of an event of type *B* is a causally necessary condition for the occurrence of a slightly later event of type *E*.

Finally, let us suppose – as is perfectly compatible with the preceding assumptions – that the relative numbers of all possible combinations of events of types *A*, *B* and *E*, throughout the whole history of the universe, are given by the following table:

	E		*Not*-E	
	A	*Not-A*	*A*	*Not-A*
B	1	0	18	12
Not-*B*	0	0	42	28

From this table, one can see that Prob(A) = 61/101, or about 0.604, while Prob(A/B) = 19/31, or slightly less than 0.613, so that, if the absolute numbers are not too large, there will be nothing especially remarkable about the fact that Prob(A/B) > Prob(A).

Next, examining the numbers that fall under '*E*', we can see that we have the following probabilities:

$$\text{Prob}(A/E) = 1;\ \text{Prob}(B/E) = 1;\ \text{Prob}(A \ \&\ B/E) = 1$$

Hence the following is true:

$$(1)\ \text{Prob}(A \ \&\ B/E) = \text{Prob}(A/E) \times \text{Prob}(B/E)$$

Similarly, examining the numbers that fall under 'not-*E*', we can see that we have the following probabilities:

$$\text{Prob}(A/\text{not-}E) = 60/100 = 0.6;\ \text{Prob}(B/\text{not-}E) = 30/100 = 0.3;$$
$$\text{Prob}(A \ \&\ B/\text{not-}E) = 18/100 = 0.18$$

So the following three equations are also true:

(2) Prob(A & B/not-E) = Prob(A/not-E) × Prob(B/not-E)

(3) Prob(A/E) > Prob(A/not-E)

(4) Prob(B/E) > Prob(B/not-E)

Hence, in a universe of the sort just described, the three types of events A, B and E form a conjunctive fork. Moreover, since there is, by hypothesis, no type of event, C, that is a common cause of events of types A and B, it is therefore the case that A, B and E constitute an open fork. This open fork then defines the relevant direction of causation as the direction that runs from events of type E towards events of the two types, A and B, that are causally necessary conditions for the occurrence of an event of type E.

In short, not only is it logically possible to have an open fork that involves a common effect, rather than a common cause, but there is no significant unlikelihood associated with the occurrence of such an open fork. The direction of open forks cannot, therefore, serve to define the direction of causation.

3.2.2 Underived laws of co-existence

John Stuart Mill suggested that, in addition, to causal laws, there could be basic laws of necessary co-existence that related simultaneous states of affairs. Are such laws possible? If one considers some candidates that might be proposed, it may be tempting, I think, to be attracted to the idea that although there can be laws of necessary co-existence, all such laws are derived, rather than basic, though this idea is far from unproblematic. Thus, consider, for example, a Newtonian world, and Newton's Third Law of Motion – that if one body, X, exerts a certain force, F, on another body, Y, then Y exerts a force equal in magnitude to F, and opposite in direction, upon X. This certainly asserts the existence of a necessary connection between simultaneous states of affairs, but is it correctly viewed as a basic law, in a Newtonian universe? Doubts arise, I think, in view of the fact that the fundamental force laws entail conclusions such as the following:

(a) It is a law that for any objects X and Y, and any time t, if X exerts a gravitational force, F, on Y at time t, then Y exerts a gravitational force, $-F$, on X at time t.
(b) It is a law that for any objects X and Y, and any time t, if X exerts an electrostatic force, F, on Y at time t, then Y exerts a gravitational force, $-F$, on X at time t.
(c) It is a law that for any objects X and Y, and any time t, if X exerts a magnetic force, F, on Y at time t, then Y exerts a gravitational force, $-F$, on X at time t.

So non-causal laws that are special instances of Newton's Third Law of Motion can be derived from the fundamental force laws, and the latter are, if one treats forces realistically, causal laws.

But this is not, of course, a derivation of Newton's Third Law of Motion itself. To have the latter, it would have to be the case that there was a law to the effect that there were only certain types of forces: gravitational, electrostatic, magnetic, and so on. Moreover, even if the latter were a law in a Newtonian universe, it would not be a causal law, and so one would still not have a derivation of Newton's Third Law of Motion from causal laws alone.

It is, accordingly, far from clear that Newton's Third Law of Motion can be derived from causal laws. But a philosopher who wishes to maintain that all basic laws are causal laws has a different response available – namely that, in a Newtonian universe, Newton's Third Law of Motion would not really be a law in the strict sense: it would be, instead, a generalization based upon the forces and force laws that have been discovered to this point.

Whether or not this response is ultimately correct, I do think that it shows at least that it is unclear whether Newton's Third Law of Motion would be a case of a basic, non-causal law of co-existence. But even if this particular example is doubtful, how can one rule out the possibility of there being such laws? Why could it not be a law, for example, that all particles with mass M have charge C, and vice versa, without that law's being derivable from any other laws whatsoever? The claim that this is not possible surely requires an argument. But what could the argument be?

In the absence of a proof of the impossibility of basic, non-causal laws of co-existence, it seems to me that one is justified in holding that such laws are logically possible. But if this is right, then Reichenbach's Principle of the Common Cause is unsound, since the extremely improbable coincidence that all particles with mass M have charge C, and vice versa, rather than being explained causally, might simply obtain in virtue of a basic, non-causal law.

3.2.3 Underived laws of co-existence, and non-accidental, open forks involving common effects

If there can be such laws, this also allows one to show that there can be open forks involving common effects that, rather than depending upon accidents of distribution, arise simply in virtue of certain laws. In particular, consider a world in which the following things are the case:

(a) The occurrence of an event of type *A* is a causally necessary condition for the occurrence of a slightly later event of type *E*.
(b) The occurrence of an event of type *B* is a causally necessary condition for the occurrence of a slightly later event of type *E*.
(c) The co-occurrence of an event of type *A* and an event of type *B* is a causally sufficient condition for the occurrence of a slightly later event of type *E*.

(d) It is a basic, non-causal law that an event of type A is always accompanied by an event of type B, and vice versa.

Then, provided that there is at least one occurrence of an event of type E, the following probabilities must obtain in virtue of (a) through (d):

$$\text{Prob}(A/E) = 1;\ \text{Prob}(B/E) = 1;\ \text{Prob}(A \ \&\ B/E) = 1$$

$$\text{Prob}(A/\text{not-}E) = 0;\ \text{Prob}(B/\text{not-}E) = 0;\ \text{Prob}(A \ \&\ B/\text{not-}E) = 0$$

It then follows that the following four equations are all true:

(1) $\text{Prob}(A \ \&\ B/E) = \text{Prob}(A/E) \times \text{Prob}(B/E)$

(2) $\text{Prob}(A \ \&\ B/\text{not-}E) = \text{Prob}(A/\text{not-E}) \times \text{Prob}(B/\text{not-}E)$

(3) $\text{Prob}(A/E) > \text{Prob}(A/\text{not-}E)$

(4) $\text{Prob}(B/E) > \text{Prob}(B/\text{not-}E)$

So the conclusion, accordingly, is that if there can be basic laws of co-existence, then there can be cases of open forks involving common effects that obtain, not by accident, but in virtue of laws of nature.

3.2.4 Simple, deterministic, temporally symmetric worlds

The next objection to the present probabilistic analysis of causation applies to any reductionist account of a Humean sort, and the basic idea is this. On the one hand, the actual world is a complex one, with a number of features that might be invoked as the basis of a reductionist account of the direction of causation. For, first of all, the direction of increase in entropy is the same in the vast majority of isolated or quasi-isolated systems (Reichenbach 1956: 117–43; Grünbaum 1973: 254–64). Second, the temporal direction in which order is propagated – such as by the circular waves that result when a stone strikes a pond, or by the spherical wave fronts associated with a point source of light – is invariably the same (Popper 1956: 538). Third, it is also a fact that all, or virtually all, open forks are open in the same direction – namely, towards the future (Reichenbach 1956: 161–3; Salmon 1978: 696).

On the other hand, causal worlds that are much simpler than our own, and that lack such features, are surely possible. In particular, consider a world that contains only a single particle, or a world that contains no fields, and nothing material except for two spheres connected by a rod that rotate endlessly about one another, on circular trajectories, in accordance with the laws of Newtonian physics. In the first world, there are causal connections between the temporal parts of the single particle. In the second world, each sphere will undergo acceleration of a constant magnitude, due to the force exerted on it by the connecting rod. So both worlds certainly contain causal relations. But both worlds are also utterly devoid of changes of entropy, of propagation of order, and of all causal forks, open or otherwise. The probabilistic analysis that we are considering, however, defines the direction of causation in terms of open forks. Simple worlds such as those just mentioned show, therefore, that this probabilistic analysis cannot be sound.

But what if the advocate of such an analysis responded by challenging the claim that such worlds contain causation? In the case of the rotating spheres' world, this could only be done by holding that it is logically impossible for Newton's Second Law of Motion to be a causal law, while in the case of the single particle world, one would have to hold that identity over time is not logically supervenient upon causal relations between temporal parts. But both of theses claims, surely, are very implausible.

In addition, however, such a challenge would also involve a rejection of the following principle:

> *The Intrinsicness of Causation in a Deterministic World*
> If C_1 is a process in world W_1, and C_2 a process in world W_2, and if C_1 and C_2 are qualitatively identical, and if W_1 and W_2 are deterministic worlds with exactly the same laws of nature, then C_1 is a causal process if and only if C_2 is a causal process.

For consider a world that differs from the world with the two rotating spheres by having additional objects that enter into causal interactions, and one of which collides with one of the spheres at some time *t*. In that world, the process of the spheres rotating around one another during some interval when no object is colliding with them will be a causal process. But then, by the above principle, the rotation of the spheres about one another, during an interval of the same length, in the simple universe, must also be a causal process.

But is the Principle of the Intrinsicness of Causation in a Deterministic World correct? Some philosophers have claimed that it is not. In particular, it has been thought that a type of causal situation to which Jonathan Schaffer (2000a: 165–81) has drawn attention – cases of 'trumping preemption' – show that the above principle must be rejected.

Here is a slight variant on a case described by Schaffer. Imagine a magical world where, first of all, spells can bring about their effects via direct action at a temporal distance, and second, earlier spells prevail over later ones. At noon, Merlin casts a spell to turn a certain prince into a frog at midnight – a spell that is not preceded by any earlier, relevant spells. A bit later, Morgana also casts a spell to turn the same prince into a frog at midnight. Schaffer argues, in a detailed and convincing way, that the simplest hypothesis concerning the relevant laws entails that the prince's turning into a frog is not a case of causal overdetermination: it is a case of preemption.

It differs, however, from more familiar cases of preemption, where one causal process preempts another by preventing the occurrence of some event that is crucial to the other process. For in this action-at-a-temporal-distance case, both processes are fully present, since they consist simply of the casting of a spell plus the prince's turning into a frog at midnight.

A number of philosophers, including David Lewis (2000), have thought that the possibility of trumping preemption shows that the Principle of the Intrinsicness of Causation in a Deterministic World is false, the idea being that there could be two qualitatively identical processes, one of which is causal and the other not. For example, at time t_1, Morgana casts a spell that a person will turn into a frog in one hour's time at a certain location. That person does turn into a frog, because there was no earlier, relevant spell. At time t_2, Morgana casts precisely the same type of spell. The person in question does turn into a frog, but the cause of this was not Morgana's latter spell, but an earlier, preempting spell.

Is this a counterexample to the Intrinsicness Principle? The answer is that it is not. Causes are states of affairs, and the state of affairs that, in the t_1 case, causes the person to turn into a frog is not simply Morgana's casting of the spell: it is the state of affairs together with the absence of earlier, relevant spells. So when the complete state of affairs that is the cause is focused upon, the two spell-casting cases are not qualitatively identical. Trumping preemption is therefore not a counterexample to the Principle of the Intrinsicness of Causation in a Deterministic World.

3.2.5 Simple, probabilistic, temporally non-symmetric worlds

The two simple, possible worlds mentioned in the preceding section were deterministic worlds, and they were also worlds that, as regards non-causal states of affairs, were precisely the same in both temporal directions. Because of the latter property, they are counterexamples to any Humean, reductionist analysis of causation. For given the complete temporal symmetry, there cannot be any Humean feature that will serve to pick out one of the two temporal directions as the direction of causation.

That complete temporal symmetry also meant, however, that there is no evidence in such worlds as to what the direction of causation is, and, for those with verificationist tendencies, this will be viewed as a reason for denying that there is any direction to causation in those worlds. What I now want to do, accordingly, is to show that there are other simple worlds that are equally counterexamples to Humean, reductionist analyses of causation, but that are not temporally symmetric, and that, because of the precise way in which they are asymmetric, are worlds that contain very strong evidence concerning the likely direction of causation.

Consider a world that contains states of affairs of types $S_0(x, t)$, $S_1(x, t)$, $S_2(x, t)$, $S_3(x, t)$, … $S_n(x, t)$, which are as follows. First, $S_0(x, t)$ is a state of affair in which absolutely nothing exists at location x at time t. Second, if i is odd, $S_i(x, t)$ consists of 2^i atomic elements of type A that are equally spaced on a circle of radius r, while if i is even, $S_i(x, t)$ consists of 2^i atomic elements of type B that are equally spaced on a circle of radius r. So, leaving aside the circular arrangement of elements, the first few states of affairs are as follows:

$S_0(x, t)$: Nothing at all

$S_1(x, t)$: *A A*

$S_2(x, t)$: *B B B B*

$S_3(x, t)$: *A A A A A A A A*

$S_4(x, t)$: *B B B B B B B B B B B B B B B B*

$S_5(x, t)$: *A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A*

Consider, now, the following two possible laws:

L_1: For every region x, and every time t, if there is a state of affairs of type $S_i(x, t)$, where i is greater than 0 and less than n, that state of affairs will continue to exist until it has existed for a temporal interval of length d, at which point it will be replaced by a state of affairs of type $S_{i+1}(x, t^*)$, where the spatial orientation of the latter state of affairs with respect to that of the temporally preceding one is completely random, while, if there is a state of affairs of type $S_n(x, t)$, that state of affairs will continue to exist until it has existed for a temporal interval of length d, at which point it will be replaced by a state of affairs of type $S_0(x, t^*)$.

L_2: For every region x, and every time t, if there is a state of affairs of type $S_i(x, t)$, where i is greater than 0, that state of affairs will continue to exist until it has existed for a temporal interval of length d, at which point it will be replaced by a state of affairs of type $S_{i-1}(x, t^*)$, where the spatial orientation of the latter state of affairs with respect to that of the temporally preceding one is completely random.

Why have I specified that it is a completely random matter how successive states of affairs are spatially oriented relative to one another? The answer is that this has been done to make it impossible, given the present account of causation, for there to be any causal forks in any world whose only law is either L_1 or L_2. For consider the transition from S_1 to S_2. If the relative spatial orientation of S_1 and S_2 is a random matter, then there is nothing that can make it the case, given the present account, that one of the two A elements in the state of type S_1 is causally related to two specific B elements in the succeeding S_2 state. All that one will be able to say is that the one total state of affairs causes the other total state of affairs, and because one cannot break this down into relations between parts of one and parts of the other, no causal forks will exist.

Suppose now that T_1 and T_2 are two types of worlds, each with the same, very large number of spatial locations. Suppose, further, that L_1 is the only law in worlds of type T_1, and that L_2 the only law in worlds of type T_2, and that in worlds of type T_1, a state of affairs of type $S_1(x, t)$ sometimes pops into existence, completely uncaused, in vacant regions of sufficient size, while, in worlds of type T_2, a state of affairs of type $S_n(x, t)$ sometimes pops into existence, completely uncaused, in vacant regions of sufficient size.

Suppose, finally, that W is a world that is either of type T_1 or of type T_2. As we have seen, because it is a completely random matter how successive states of affairs are spatially oriented relative to one another, there cannot be, given the probabilistic analysis of causation that we are now considering, any forks in world W – and, a fortiori, any open forks. It therefore follows, on this analysis of causation, that there is no direction of causation, and so no causation, in world W.

But this conclusion is unsound. The information that one has about the world makes it very likely that there is causation in world W, and that it has a certain direction. For compare worlds of type T_1 with worlds of type T_2. In worlds of the former sort, the only type of state of affairs that comes into existence uncaused is a state of affairs of type $S_1(x, t)$, and since this consists of only two atomic elements of type A, it is not especially unlikely that such a state of affairs should come into existence uncaused. By contrast, in worlds of type T_2, the type of state of affairs that comes into existence uncaused is a state of affairs of type $S_{\mathrm{n}}(x, t)$, and since this may very well consist of an enormous number of atomic elements – since n can be any number one wants, such as 10^{100} – all of them of the same type, equally spaced on a circle, it may, by contrast, be extraordinarily unlikely that such a state of affairs should come into existence uncaused.

The upshot, in short, is that given a world W that is either of type T_1 or of type T_2, it is much more likely that W is of type T_1 than of type T_2, and so it is much more likely that the direction of causation runs from states of affairs of type $S_1(x, t)$ to type $S_2(x, t)$ and on to type $S_n(x, t)$ than that it runs in the opposite direction.

Finally, though worlds of types T_1 and T_2 do involve laws that are not completely probabilistic, since the temporal interval at which one state of affairs is replaced by another is fixed, that it not essential, and one could replace laws L_1 and L_2 by totally probabilistic laws in which each of the relevant states of affairs has a certain half-life, so that there would merely be a certain probability that a given state of affairs would, within a given temporal interval, be replaced by the next state in the relevant order. The resulting world types – T_1* and T_2* – would then be completely probabilistic worlds, but that would not alter the fact that it would be much more likely that the direction of causation was from states of affairs of type $S_1(x, t)$ to states of affairs of type $S_2(x, t)$ and on to states of affairs of type $S_{\mathrm{n}}(x, t)$, rather than in the opposite direction.

The conclusion, accordingly, is that there are simple, probabilistic worlds in which causation is present, and in which there is good reason for viewing one of the two possible temporal directions as the direction of causation, but where the probabilistic analysis of causation that we are considering mistakenly entails that no causation is present.

3.2.6 Temporally 'inverted', twin universes

It is the year 4004 BC. A Laplacean-style deity is about to create a world rather similar to ours, but one where Newtonian physics is true. Having selected the year AD 3000 as a good time for Armageddon, the deity works out what the world will be like at that point, down to the last detail. He then creates two spatially unrelated worlds: the one just mentioned, together with another whose initial state is a flipped-over version of the state of the first world immediately prior to Armageddon – in other words, the two states agree exactly, except that the velocities of the particles in the one state are exactly opposite to those in the other.

Consider, now, any two complete temporal slices of the first world, A and B, where A is earlier than B. Since the worlds are Newtonian ones, and since the laws of Newtonian physics are invariant with respect to time reversal, the world that starts off from the reversed, AD 3000 type of state will go through corresponding states, B^* and A^*, where these are flipped-over versions of B and A respectively, and where B^* is earlier than A^*.

So while the one world goes from a 4004 BC, Garden of Eden state to an AD 3000, pre-Armageddon state, the other world will move from a reversed, pre-Armageddon type of state to a reversed, Garden of Eden type of state.

In the first world, the direction of causation will coincide with such things as the direction of increase in entropy, the direction of the propagation of order in non-entropically irreversible processes, and the direction defined by most open forks. But in the second world, where the direction of causation runs from the initial state created by the deity – that is, the flipped-over AD 3000 type of state – through to the flipped-over 4004 BC type of state, the direction in which entropy increases, the direction in which order is propagated, and the direction defined by open forks will all be the opposite one. So if any of the latter is used to define the direction of causation, it will generate the wrong result in the case of the second world. The probabilistic analysis of causation that we are presently considering assigns, therefore, the wrong direction to causation in the case of the second world.

3.2.7 Causally ambiguous situations in probabilistic worlds

A reductionist analysis of causation in terms of relative frequencies is also exposed to a variety of 'underdetermination' objections, the thrust of which is that fixing all of the non-causal properties of, and relations between, events, including all relative frequencies, does not always suffice to fix what causal relations there are between events. Indeed, the arguments in question support much stronger conclusions – such as, for example, the conclusion that even if one also fixes what laws there are, both causal and non-causal, along with the direction of causation for all possible causal relations that might obtain, that still does not suffice to settle what causal relations there are between events.

One such argument can be set out as follows.[2] First, one needs to ask whether statements of causal laws can involve the concept of causation. Consider, for example, the following statement: 'It is a law that for any object x, the state of affairs that consists of x's having property F causes a state of affairs that consists of x's having property G.' Is this an acceptable way of formulating a possible causal law?

Some philosophers contend that it is not, and that the correct formulation is, instead, along the following lines:

(*) 'It is a causal law that for any object x, if x has property F at time t, then x has property G at $(t + \Delta t)$.'

But what reason is there for thinking that it is the latter type of formulation that is correct? Certainly, as regards intuitions, there is no reason why there should not be laws that themselves involve the relation of causation. But in addition, the above claim is open to the following objection. First, the following two statements are logically equivalent:

(1) For any object x, if x has property F at time t, then x has property G at $(t + \Delta t)$.
(2) For any object x, if x lacks property G at time $(t + \Delta t)$, then x lacks property F at t.

Now replace the occurrence of (1) in (*) by an occurrence of (2), so that one has:

(**) 'It is a causal law that for any object x, if x lacks property G at time $(t + \Delta t)$, then x lacks property F at time t.'

The problem now is that it may very well be the case that while (*) is true, (**) is false, since its being a causal law that for any object x, if x has property F at time t, then x has property G at $(t + \Delta t)$. G certainly does not entail that there is a backward causal law to the effect that for any object x, if x lacks property G at time $(t + \Delta t)$, then x lacks property F at t. So anyone who holds that (*) is the correct way to formulate causal laws needs to explain why substitution of logically equivalent statements in the relevant context does not preserve truth.

By contrast, no such problem arises if one holds that causal laws can instead be formulated as follows:

It is a law that for any object x, the state of affairs that consists of x's having property F at time t causes a state of affairs that consists of x's having property G at time $(t + \Delta t)$.

Let us assume, then, that the natural way of formulating causal laws is acceptable. The next step in the argument involves the assumption that probabilistic laws are logically possible. Given these two assumptions, the following presumably expresses a possible causal law:

L_1: It is a law that, for any object x, x's having property P for a time interval Δt causally brings it about, with probability 0.75, that x has property Q.

The final crucial assumption is that it is logically possible for there to be uncaused events.

Given these assumptions, consider a world, *W*, where objects that have property P for a time interval Δt go on to acquire property Q 76% of the time, rather than 75% of the time, and where this occurs even over the long term. Other things being equal, this would be grounds for thinking that the relevant law was not L_1, but rather:

L_2: It is a law that, for any object x, x's having property P for a time interval Δt causally brings it about, with probability 0.76, that x has property Q.

But other things might not be equal. In the first place, it might be the case that L_1 was derivable from a very powerful, simple and well-confirmed theory, whereas L_2 was not. Second, one might have excellent evidence that there were totally uncaused events involving objects' acquiring property Q, and that the frequency with which that happened was precisely such as would lead to the expectation, given law L_1, that situations in which an object had property P for a time interval Δt would be followed by the object's acquiring property Q 76% of the time.

If that were the case, one would have reason for believing that, on average, over the long term, of the 76 cases out of a 100 where an object that has property P for Δt and then acquires property Q, 75 of those cases will be ones where the acquisition of property Q is caused by the possession of property P, while one out of the 76 will be a case where property Q is spontaneously acquired.

There can, in short, be situations where there would be good reason for believing that not all cases where an object has property P for an interval Δt, and then acquires Q, are causally the same. There is, however, no hope of making sense of this given a reductionist analysis of causation in terms of relative frequencies. For the cases do not differ with respect to any non-causal properties and relations, including relative frequencies, nor with respect to causal or non-causal laws, nor with respect to the direction of causation in any potential causal relations. So the present approach is unable to deal with such causally ambiguous, probabilistic situations.

3.2.8 Causation without increase in probability

We have not yet considered the most fundamental claim involved not only in the attempt to analyse causation in terms of relative frequencies, but, indeed, in all probabilistic analyses of causation – the proposition, namely, that causes always make their effects more likely, in some appropriate sense. Is this claim true? The answer appears to be that it is not, as even some philosophers who are sympathetic to the general idea that there is *some* connection between causation and probability – such as Daniel Hausman (1998) – have realized. For consider the following.

Assume that there are atoms of type *T* that satisfy the following conditions:

(1) Any atom of type *T* must be in one of the three mutually exclusive states – *A*, *B* or *C*.
(2) The probabilities that an atom of type *T* in states *A*, *B* and *C*, respectively, will emit an electron are, respectively, 0.9, 0.7 and 0.2.
(3) The probabilities that an atom of type *T* is in state *A* is 0.5; in state *B*, 0.4; and in state *C*, 0.1.

Now, given that, for example, putting an atom of type *T* into state *B* would be quite an effective means of getting it to emit an electron, it is surely true that, if it is in state *B*, and emits an electron, then its being in state *B* is a probabilistic cause of its emitting an electron. But this would not be so if the above account were correct. For if *D* is the property of emitting an electron, the unconditional probability that an atom of type *T* will emit an electron is given by Prob(*D*) = Prob(*D*/*A*) × Prob(*A*) + Prob(*D*/*B*) × Prob(*B*) + Prob(*D*/*C*) × Prob(*C*) = (0.9)(0.5) + (0.7)(0.4) + (0.2)(0.1) = 0.75. But the conditional probability of *D* given *B* was specified as 0.7. We have, therefore, that Prob(*D*) > Prob(*D*/*B*). So if a cause had to raise the probability of its effect, it would follow that an atom of type *T*'s being in state *B* could not be a probabilistic cause of its emitting an electron. This, however, is unacceptable. So the thesis that a cause must raise the probability of its effect, in the relevant sense, must be rejected.

The thesis that causes necessarily make their effects more likely is exposed, therefore, to a decisive objection. The basis of this objection is the possibility of there being one or more other causal factors that are incompatible with the given factor, and more efficacious than it. For, given such a possibility, events of type *C* may be the cause of events of type *E* even though the probability of an event of type *E*, given the occurrence of an event of type *C*, is less than the unconditional probability of an event of type *E*.

But is there nothing, then, in the rather widely shared intuition that causation is related to increase in probability? The answer is that causation may be related to increase in probability, but not in the way proposed by those who favour a probabilistic analysis of causation. What this other way is will emerge in Section 5. The crucial point for present purposes, however, is that the relation in question cannot be used as part of a probabilistic *analysis* of causation, since the relation itself turns out to involve the concept of causation.

4 Causation and objective chances

First, however, there is another very important reductionist alternative that needs to be considered: the idea, namely, that causation can be analysed in terms of objective chances, together with non-causal states of affairs.

If this programme is to be carried out, one needs to hold that objective chances are not to be analysed in terms of causation, and here there are two main possibilities. The first, and by far the more common, view is that objective chances are themselves ontologically ultimate states of affairs, and so, a fortiori, not analysable in terms of causation. The other, and much less commonly adopted, view is that objective chances, rather than being either ontologically ultimate or analysable in causal terms, supervene upon laws, characterized non-causally, together with non-causal states of affairs.

4.1 Causation and ontologically ultimate, objective chances

A number of philosophers – such as Edward Madden and Rom Harré (1975), Nancy Cartwright (1989), and C. B. Martin (1993) – have advocated both an ontology in which irreducible dispositional properties, powers, propensities, chances and the like occupy a central place, and maintained that such an ontology is relevant to causation. Often, however, the details have been rather sparse. But a clear account of the basic idea of analysing causation in terms of objective chances was set out in 1986 both by D. H. Mellor and by David Lewis (1986) and then, more recently, Mellor has offered a very detailed statement and defence of this general approach in his book *The Facts of Causation* (1995).

4.1.1 Lewis's account: counterfactuals and objective chances

This general approach to causation was briefly sketched by David Lewis in a post-script to his article 'Causation':

> there is a second case to be considered: *c* occurs, *e* has some chance *x* of occurring, and as it happens *e* does occur; if *c* had not occurred, *e* would still have had some chance *y* of occurring, but only a very slight chance since *y* would have been very much less than *x*. We cannot quite say that without the cause, the effect would not have occurred; but we can say that without the cause, the effect would have been very much less probable than it actually was. In this case also, I think we should say that *e* depends causally on *c*, and that *c* is a cause of *e*.
>
> (Lewis 1986: 176)

Lewis advanced this as an account of probabilistic causation. But, as Lewis notes, by employing chances where the probabilities are exactly one and exactly zero – as contrasted with infinitesimally close to one and zero – one can view this as a general account of causation that covers non-probabilistic causation as well as probabilistic causation.

A feature of this account that does not seem especially plausible is the requirement that, in the absence of the cause, the probability of the effect would have been *much* lower. If one drops that requirement, Lewis's account is as follows:

(1) An event c causes an event e if and only if there is a chain of causally dependent events linking e with c.
(2) An event e is causally dependent upon an event c if and only if there are numbers x and y such that (a) if c were to occur, the chance of e occurring would be equal to x; (b) if c were not to occur, the chance of e occurring would be equal to y; and (c) x is greater than y.

4.1.2 Mellor's account of causation: objective chances and strong laws

A very closely related analysis was set out by D. H. Mellor in his book *The Facts of Causation*, but Mellor's account is much more detailed and wide-ranging, and he offers a host of arguments in support of the central aspects of the analysis, including the crucial claim that a cause must raise the probability of its effect. Mellor also diverges from Lewis in rejecting a regularity account of laws in favour of a view according to which even basic laws of nature can exist without having instances.

Mellor's approach, in brief, is as follows. First, Mellor embraces an ontology involving objective chances, where the latter are ultimate properties of states of affairs, rather than being logically supervenient upon causal laws together with non-dispositional properties, plus relations. Second, Mellor proposes that chances can be defined as properties that satisfy three conditions: (1) The Necessity Condition: if the chance of P's obtaining is equal to one, then P is the case; (2) The Evidence Condition: if one's total evidence concerning P is that the chance of P is equal to k, then one's subjective probability that P is the case should be equal to k; (3) The Frequency Condition: the chance that P is the case is related to the corresponding relative frequency in the limit.[3] Third, chances enter into basic laws of nature. Fourth, Mellor holds that even basic laws of nature need not have instances, thereby rejecting reductionist accounts in favour of a realist view. Fifth, any chance that P is the case must be a property of a state of affairs that temporally precedes the time at which P exists, or would exist. Finally, and as a very rough approximation, a state of affairs C causes a state of affairs E if and only if there are numbers x and y such that (1) the total state of affairs that exists at the time of C – including laws of nature – entails that the chance of E is x, (2) the total state of affairs that would exist at the time of C, if C did not exist, entails that the chance of E is y, and (3) x is greater than y.[4]

4.2 Objections

Objections to this approach to causation are of three main types. First, this approach employs the Stalnaker–Lewis style of counterfactuals, and it can be objected that such a closest-worlds account of counterfactuals is unsound. Second, there are objections that are directed against the view that objective chances are ontologically ultimate properties. Third, there are objections to the effect that, even given this view of objective chances, the resulting account of causation is unsound.

4.2.1 Closest-worlds conditionals

The first objection is that an analysis of counterfactuals in terms of similarities across possible worlds is exposed to a number of serious objections. One of the most important is a type of objection originally advanced by philosophers such as Jonathan Bennett (1974) and Kit Fine (1975), who contended that a Stalnaker–Lewis account generates the wrong truth values for counterfactuals in which the consequent could only be true if the world were radically different from the actual world. Thus Fine, for example, argued that the following counterfactual would turn out to be false on a Stalnaker–Lewis approach:

> If Nixon had pressed the button, there would have been a nuclear holocaust.

In response to this objection, David Lewis, in his article 'Counterfactual Dependence and Time's Arrow' (1979), argued that by assigning certain weights to big miracles, to perfect matches of particular facts throughout a stretch of time, and to small miracles, one could make it the case that the Nixon-and-the-button counterfactual came out true, rather than false. Lewis's escape, however, cannot handle the general problem that Fine, Bennett, and others, raised. For Lewis's solution depends upon the fact that Nixon's pressing the button is an event which would have multiple effects, and which thus is such that it would require a very big miracle to remove all traces of that event, and so achieve a perfect match with the future of the actual world. As a result, one needs merely to construct a case involving an event that has only a single effect. This is easily done, and then Lewis's account of similarity does not block the counterexample (Tooley 2003).

So the use of closest-worlds counterfactuals is not satisfactory. However, one needs to ask whether the use of such conditionals is an essential feature of any analysis of causation in terms of objective chances. Initially, it might seem that it is. For the analysis must refer not just to the chance, at the time of the cause *C*, of the effect *E*, but also to the chance that *E would* have occurred if *C* had not occurred. Accordingly, counterfactual conditionals are certainly needed, and in the context of giving an analysis of causation, one cannot, of course, adopt a causal account of the truth conditions of counterfactuals. So what alternative is there to a closest-worlds account?

The answer is that there is another alternative – namely, one that arises out of the idea that the chances that exist at a given time, rather than supervening on categorical states of affairs that exist at that time together with probabilistic causal laws, supervene instead upon categorical states of affairs together with non-probabilistic, non-causal laws linking categorical properties at a time to chances at that time. For if this view can be defended, then rather than asking about the chance that *E* would occur in the closest worlds where *C* does not occur, one can ignore past and future similarities, and ask instead about the chance that *E* would occur in those worlds where *C* does not occur and that are *most similar at the time of C* to the world where *C* occurs.

The idea, in short, is that one can shift from closest-worlds counterfactuals to closest-momentary-slices counterfactuals, thereby avoiding the objections to which the former are exposed.

4.2.2 Logical connections between temporally distinct states of affairs

The next four objections are directed against the view that objective chances are ontologically ultimate properties of things at a time. First, the postulation of objective chances, understood as intrinsic properties of things, involves the postulation of non-Humean states of affairs, since objective chances, thus understood, enter into logical relations with distinct states of affairs. It is true that those logical relations will, in general, be probabilifying ones, rather than relations of logical entailment, and one might try to argue that while the latter are problematic, the former are not. That line of argument, however, seems to me very dubious. But even if it could be sustained, it would not answer the present objection. For an account of objective chances must also cover the limiting case where the probability in question, rather than being at most infinitesimally close to one, is precisely one.

Consider, for example, the law of conservation of charge, and suppose that the universe contains, at time t, a total net charge of n units. On the present account, objective chances must be present at time t that logically entail that the total net charge of the universe at any later time ($t + \Delta t$) is also equal to n.

David Hume contended that it is logically impossible for there to be logical connections between distinct states of affairs, and this thesis is, I think, very widely accepted today. Thus Bas van Fraassen (1989), for example, views it as a decisive objection to various realist conceptions of laws of nature. For if laws of nature are conceived of as second-order relations between universals – as by Dretske (1977), Tooley (1977, 1987), and Armstrong (1983) – or as structureless states of affairs – as by Carroll (1994) – they have to be identified via the fact that they are states of affairs that entail the existence of corresponding cosmic regularities involving only Humean properties – and so it appears that laws of nature, thus interpreted, entail Humean states of affairs which they neither are identical with nor overlap. So it *appears* that one has logical relations between ontologically distinct states of affairs.

It turns out, however, that whether laws of nature, thus conceived, do involve non-Humean states of affairs depends upon precisely what account is given of the ontology involved, since it can be shown that if transcendent universals are admitted, there are metaphysical hypotheses concerning the existence of such universals that do clearly and straightforwardly entail the existence of corresponding regularities (Tooley 1987: 123–9) – the basic idea being that if only certain transcendent universals exist, this must limit what states of affairs can exist at the level of particulars, and it will do so without introducing any logical relations between distinct states of affairs.

By contrast, when objective chances are conceived of as intrinsic properties of things at a time, the existence of such properties surely does entail, at least in the limiting cases, the existence of logical connections between distinct states of affairs, since one has a logical entailment between things' having intrinsic properties at one time, and things' having intrinsic properties at other times. Accordingly, if Hume's thesis is correct, we have here a decisive objection to the present account of objective chances.

4.2.3 Basic laws

The first objection leads immediately to a second, which is concerned with the implications that the view that objective chances are ontologically ultimate has with regard to the nature of basic laws. Consider, for example, a Newtonian world. One normally thinks that, in such a world, Newton's Second Law of Motion – F = MA – is a basic law that relates the mass of an object at a given time, and the force acting on it at that time, to its acceleration at a later time. (Because time is dense, and there is no next moment, a somewhat more complex formulation in terms of intervals is needed here. But we can ignore that, as it does not affect the present point.) Suppose, however, that there are ontologically ultimate, objective chances, and that causation is to be analysed in terms of them. Then we need to think of the relation between force and mass at one time, and acceleration at a later time, in a different way. For what one then has are two connections:

(1) There is a basic law of nature that connects up things existing at one and the same time – namely, on the one hand, force and mass and, on the other hand, an objective chance equal to one of a later acceleration equal to F/M.
(2) There is a logically necessary connection between an objective chance that exists at one time – of the object's undergoing a later acceleration equal to F/M – and the acceleration of the object at that later time.

So rather than having a causal law connecting states of affairs existing at different times, what we have is a law of dependence connecting something existing at one time – namely, a certain objective chance – with other things existing at the *very same* time – namely, an object's having a certain mass, and being acted upon by a certain force. The only laws that there are, accordingly, if causation is analysed in terms of ontologically ultimate objective chances are laws connecting simultaneous states of affairs, and connections between states of affairs existing at different times, rather than being underwritten by laws of nature, are logically necessary, if the world is deterministic.

As was mentioned earlier, some philosophers have held that there can be both basic laws of co-existence, and basic laws connecting things at different times, while other philosophers have been suspicious of the idea of basic laws of co-existence, and have favoured the view that all laws of co-existence are derived from basic causal laws. The argument for the latter view is unclear. Nevertheless, I think that one can see, at least dimly, why one might find basic laws of co-existence somehow less intelligible, or more problematic, than basic causal laws. By contrast, the opposite view seems to have no evident appeal at all. For if there can be basic laws of nature that link together things at one and the same time, why should there be any problem with basic laws of nature that link together states of affairs at different times?

The upshot is that the idea of analysing causation in terms of objective chances has consequences with regard to the types of basic laws that are possible – consequences that, on the face of it, do not seem at all plausible.

4.2.4 The infinite states of affairs objection

The third objection to the view that objective chances are ontologically ultimate properties of things at a time can be put as follows. Imagine that the world is deterministic, that every temporal interval is divisible, and that all causation involves continuous processes. Suppose that x at time t has an objective chance equal to 1 of being C at time $(t + \Delta t)$. Then there are an infinite number of moments between t and $(t + \Delta t)$, and for every such moment, t^*, it must be the case either that x at time t has an objective chance equal to 1 of being C at time t^*, or that x at time t has an objective chance equal to 1 of not being C at time t^*. But then, if objective chances are ontologically ultimate, intrinsic properties of things at a time, it follows that x at time t must have an infinite number of intrinsic properties – indeed, a non-denumerably infinite number of properties.

This view of the nature of objective chances involves, accordingly, a very expansive ontology indeed. By contrast, if objective chances, rather than being ontologically basic, supervene on categorical properties plus causal laws, this infinite set of intrinsic properties of x at time t disappears, and all that one need have is a single, intrinsic, categorical property – or a small number of such properties – together with relevant laws of nature.

4.2.5 The compatibility of objective chances objection

The thrust of the fourth and final objection to the view that objective chances are ontologically ultimate is that there are pairs of objective chances that, intuitively, are perfectly compatible, but that would be logically incompatible on the present view.

The argument is as follows. Consider the following three objective chances:

(1) P = an objective chance of 0.7 of property C in Δt.
(2) Q = an objective chance of 0.2 of property D in Δt.
(3) R = an objective chance of 0.7 of property C in Δt and an objective chance of 0.2 of property D in Δt.

Clearly, something might have both property P and property Q. Suppose, then, that it is a non-causal law that anything that comes to have the categorical property A also acquires both property P and property Q at the same time. Then the probability that something that acquired property A would acquire certain combinations of properties in Δt would be as follows:

Both C and D: $(0.7)(0.2) = 0.14$
C, but not D: $(0.7)(0.8) = 0.56$
D, but not C: $(0.3)(0.2) = 0.06$
Neither C nor D: $(0.3)(0.8) = 0.24$

Propensity R, as defined above, is just a combination of propensities P and Q, and the probabilities that something with propensity R will acquire the various combinations of properties just listed would be precisely the probabilities associated with the joint possession of propensities P and Q.

Consider, now, a propensity, S, that can be described in ordinary language as follows: Propensity S gives rise either to property C or to property D, but never to both, and the probability of its giving rise to C is 0.7, while the probability of its giving rise to D is 0.2. Clearly, S is not identical with the conjunction of P and Q, nor with R, since, given S, there are different probabilities associated with the combinations of properties considered above, namely:

Both C and D: 0.0
C, but not D: 0.7
D, but not C: 0.2
Neither C nor D: 0.1

If objective chances are ontologically ultimate, how is S to be defined? The answer will depend upon precisely what the correct account is of objective chances, so understood. Earlier, I mentioned Mellor's proposed analysis. But one of its clauses involves the term 'should', and, as it seems inappropriate for a characterization of objective chances to incorporate any normative language, Mellor's account seems problematic.

The type of account that seems to me preferable can be illustrated by the following analysis of what it is to have propensity R:

x has propensity R at time t

means the same as:

There is some intrinsic property P such that, first, x has property P at time t; second, x's having property P at time t does not logically supervene upon a state of affairs that involves either the existence of certain laws of nature, causal or otherwise, or x's having some relevant categorical property, either at time t, or at any other time; and, third, the logical probability that x has property C at time t^*, given that x has property P at time t, and regardless of whatever other intrinsic properties x has at time t, is equal to 0.7, while the logical probability that x has property D at time t^*, given that x has property P at time t, and regardless of whatever other intrinsic properties x has at time t, is equal to 0.2.

With this as a model, let us now consider how the possession of propensity S is to be analysed. In the case of propensity R, probabilities are assigned to each of the two 'effect' properties – C and D. Obviously this cannot be done in the case of propensity S, since the probability that the thing in question will acquire both property C and property D is equal to zero, and this can be generated by an assignment of probabilities to each of C and D only if at least one of those probabilities is equal to zero, which is not the case.

What is needed, accordingly, is an analysis in which probabilities are assigned to at least three of the four relevant combinations of possibilities:

> x has propensity S at time t

means the same as:

> There is some intrinsic property P such that, first, x has property P at time t; second, x's having property P at time t does not logically supervene upon a state of affairs that involves either the existence of certain laws of nature, causal or otherwise, or x's having some relevant categorical property, either at time t, or at any other time; and, third, the logical probability that x has property C, but not property D, at time t^*, given that x has property P at time t, and regardless of whatever other intrinsic properties x has at time t, is equal to 0.7, while the logical probability that x has property D, but not property C, at time t^*, given that x has property P at time t, and regardless of whatever other intrinsic properties x has at time t, is equal to 0.2 and, finally, the logical probability that x has neither property C, nor property D, at time t^*, given that x has property P at time t, and regardless of whatever other intrinsic properties x has at time t, is equal to 0.1.

Consider now another propensity, T, that can be described in ordinary language as follows: Propensity T gives rise either to property C or to property D, but never to both, and the probability of its giving rise to C is 0.5, while the probability of its giving rise to D is 0.3. The crucial question now is whether an object at one and the same time could possess both property S and property T, and the answer is that this is certainly possible. For that would just mean that there would be different routes by which the object in question might acquire property C – in one case, in virtue of having property S and, in the other case, in virtue of having property T.

The problem is that the above analysis of what it is to have propensity S, together with a parallel analysis of what it is to have propensity T, entails that it is logically impossible for any object to have both of those properties at the same time. For the definition of propensity *S* entails that if something has propensity *S* at time *t*, together with any other intrinsic properties whatsoever – including propensity *T* – then the probability that *x* has property *C* at time *t** is equal to 0.7, whereas the corresponding definition of propensity *T* entails that if something has propensity *T* at time *t*, together with any other intrinsic properties whatsoever – including propensity *S* – then the probability that *x* has property *C* at time *t** is equal to 0.5.

How do things compare if objective chances, rather than being viewed as ontologically ultimate, are analysed along causal lines? To answer that question, we need to have a causal account in front of us. Such an account can easily be arrived at by generalizing upon a causal analysis of dispositional properties. So consider, for example, water-solubility. According to a familiar type of account, the statement that *x* is water-soluble is to be analysed as saying that *x* possesses some categorical property, *P*, such that there is a law of nature, *L*, that entails that, for any *y*, the state of affairs that consists of *y*'s possessing property *P* at any time *t*, and *y*'s being in water at time *t*, immediately causes *y* to dissolve.

This account of dispositional properties is easily converted into an account of objective chances. Precisely how the latter should be formulated depends upon the correct account of the logical form of probabilistic causal laws, but one natural formulation runs as follows:

> *x* at time *t* has an objective chance equal to *k* of being *C* at time *t**

means the same as:

> There is some intrinsic, categorical property, *P*, such that, first, *x* has property *P* at time *t*, and, second, there is a law of nature, *L*, to the effect that for any *y*, and any time *u*, the probability that *y*'s having property *P* at time *u* causes *y*'s having property *C* at time *u**, given that *y* has property *P* at time *u*, is equal to *k*.

The point now is that, given this type of account, something can have both propensity *S* and propensity *T* at the same time. The reason is that the probabilities that enter into the causal analysis are not probabilities, for example, that *x* will have property *C* at time *t**; they are, rather, probabilities that a certain intrinsic property of *x* at time *t* will cause *x* to have property *C* at time *t**, and there is no incompatibility involved if *x* has two intrinsic properties, *P* and *Q*, at time *t*, where the probability that possession of property *P* at time *t* will give rise to *x*'s possessing property *C* at time *t** is equal to 0.7, while the probability that possession of property *Q* at time *t* will give rise to *x*'s possessing property *C* at time *t** is equal to 0.5.

In short, there are sets of objective chances that are, intuitively, perfectly compatible, and that are compatible given a causal analysis, but that would be logically incompatible if chances were ontologically ultimate properties of a thing at a time.

4.2.6 Underdetermination objections

Suppose now that one could somehow overcome the four objections just set out against the thesis that objective chances are ontologically ultimate. There would still be at least three very strong reasons for holding that causation cannot be analysed in terms of objective chances, so understood.

First, there are underdetermination objections. For recall the argument set out in Section 3.2.7 (p. 94–6) for the conclusion that there can be situations that differ causally, even though they do not differ with respect to relevant non-causal properties and relations, nor with respect to causal or non-causal laws, nor with respect to the direction of causation in any potential causal relations. Given this conclusion, if objective chances are logically supervenient upon causal laws plus non-causal states of affairs, then the cases do not differ with respect to objective chances either. But even if one rejected the latter supervenience claim, and held that objective chances were ultimate, irreducible properties, that would not alter things, since the relevant objective chances would still be the same in both cases. The earlier argument supports, accordingly, the following, stronger conclusion that applies to any attempt to analyse causal relations in terms of objective chances: causal relations between events are not logically supervenient upon the totality of states of affairs involving non-causal properties of, and relations between, events, all of the laws, both causal and non-causal, all of the dispositional properties, propensities, and objective chances and, finally, the direction of causation for all possible causal relations that might obtain.

4.2.7 The objection to the probability-raising condition

Next, just as in the case of probabilistic accounts of causation of a Humean, reductionist sort, any analysis of causation in terms of objective chances is also exposed to the objection that causes need not raise the probability of their effects. For although it is possible, by adopting Lewis's distinction between causation and causal dependence, to argue – as Lewis does – that an analysis of causation in terms of objective chances does not entail that *causes* always raise the probabilities of their effects, the objection in question still applies, since one can show that a cause need not raise the probability of its effect even in the case of *direct* causation.

To establish that this is so, the argument that was offered earlier to show that a cause need not raise the probability of its effect needs to be modified slightly, so that, first, it deals with direct causal connections, and, second, it refers to objective chances, rather than to conditional and unconditional probabilities. This can be done as follows. Suppose that there is a type of atom, *T*, and relevant laws of nature that entail the following:

(1) Any atom of type *T* must be in one of the four mutually exclusive states – *A*, *B*, *C* or *D*.

(2) Any atom of type T in state A has an objective chance of 0.999 of moving directly into state D; an atom in state B has an objective chance of 0.99 of moving directly into state D; an atom in state C has an objective chance of 0 of moving directly into state D.
(3) There is a certain type of situation – S – such that any atom of type T in situation S must be in either state A or state B.

Suppose now that x is an atom of type T, in situation S, in state B, and which moves directly into state D. Given that, for example, shifting an atom of type T from state C into state B would be quite an effective means of getting it into state D, it is surely true that x's being in state D is probabilistically caused by x's having been in state B. But this would not be so if the above account were correct. For consider what would have been the case if x had not been in state B. Given that x was in situation S, x would, in view of (3), have been in state A. But then x's objective chance of moving directly into state D would have been 0.999, and so is higher than what it is when the atom is in state B.

The point here, as before, is that a given type of state may be causally efficacious, but not as efficacious as alternative states and, because of this, it is not true that even a direct cause need raise the probability of its effect, contrary to what is required by the above analysis.

4.2.8 Objective chances and a causal theory of time

The final objection starts out from the observation that if there is, at location s and time t, a certain objective chance of a state of affairs of type E, this is not, of course, equal to the probability that there is a state of affairs of type E somewhere in the universe: it is, rather, the probability that there is a state of affairs of type E in a location *appropriately related* to s and t.

What does this mean in the case of time? If backward causation is logically possible – as Lewis believes, and as Mellor does not – then it would seem that there could be an objective chance at location s and time t that was the chance that there is an event of type E at a certain temporal distance either before or after t. Such chances would be 'bi-directional'. But let us set those aside, and consider only the cases where a chance of there being an event of type E is either a chance of there being an event of type E at a later time, or else, a chance of there being an event of type E at an earlier time. All such chances, then, would themselves incorporate a temporal direction – either the later-than direction, or the earlier-than direction. But this means that if one proceeds to analyse causation in terms of objective chances that are not of a bi-directional sort, one cannot, on pain of circularity, analyse the direction of time in terms of the direction of causation.

Many philosophers, of course, reject a causal analysis of the direction of time, and it may be that they are right in so doing. The problem here, however, is that the impossibility of a causal theory of the direction of time would follow *immediately* from the analysis of causation, and this does not seem right, since then it would be rather puzzling why a substantial number of philosophers have been attracted to a causal theory of the direction of time.

5 Probability and a realist approach to causation

We have considered two attempts to offer a reductionist, probabilistic analysis of causation: one in terms of relative frequencies, and the other in terms of non-Humean states of affairs involving objective chances, viewed as ontologically ultimate. We have seen that both approaches are open to a large number of very strong objections and, in the light of that, it seems to me extremely unlikely that either approach is tenable.

Given this, it may well be tempting to conclude that the whole idea that some concept of probability enters into the analysis of causation is mistaken. But that conclusion would be premature at this point. For it may be that the failures of the present accounts are traceable to the fact that they are reductionist approaches. We need to consider, then, whether a satisfactory realist account of causation can be given, and one that involves some concept of probability. In this section I shall argue that that is the case.

Until relatively recently, realist approaches to causation – as advanced, for example, by Elizabeth Anscombe (1971) – almost always involved the idea that causation is directly observable and, accordingly, the related view that the concept of causation does not stand in need of any analysis: it can be viewed as analytically basic. But as I have argued elsewhere (Tooley 1990a), there are strong arguments against the view that causation is directly observable in any sense that would justify one in holding that the concept of causation is analytically basic.

If that is right, then either the concept of causation – or some other causal concept, such as that of a causal law – must be a theoretical concept, and so it is not surprising that this type of realist approach to causation has emerged only relatively recently. For serious exploration of this type of approach required, as I noted earlier, two philosophical advances – one semantical, the other epistemological. As regards the former, one needed a non-reductionist account of the meaning of theoretical terms. A paper by F. P. Ramsey written in 1929 contained the crucial idea that was needed for a solution to this problem, and the outlines of an account were then set out, albeit very briefly and almost in passing, by R. B. Braithwaite (1953: 79). It was, however, still some time before careful and generally satisfactory accounts were provided by R. M. Martin (1966) and David Lewis (1970).

As regards the epistemological issue, one needed to have reason for thinking that theoretical statements, thus interpreted, could be confirmed. That this could not be done via induction based on instantial generalization had in effect been shown by Hume (1739, Part IV, Section 2), so the question was whether there was some other legitimate form of non-deductive inference. Gradually, the idea of the method of hypothesis (hypothetico-deductive method, abduction, inference to the best explanation) emerged, and, although by no means uncontroversial, this alternative to instantial generalization is at least widely accepted by contemporary philosophers.

These two developments opened the door to the idea of treating causation as a theoretical relation, and two main accounts have now been advanced. According to the one which was advanced by David Armstrong and Adrian Heathcote (1991) and then developed in more detail by Armstrong in his book, *A World of States of Affairs* (1997, esp. 216–33), all basic laws are causal laws, so that an account of the necessitation involved in basic laws of nature *ipso facto* provides an account of causal necessitation. According to the other account, by contrast, basic laws need not be causal laws, so that the relation of causation cannot be identified with a general relation of nomic necessitation. How, then, is causation to be defined? The answer offered by the second approach is, first, that causal laws must satisfy certain postulates involving probabilistic relations and, second, that causation can then be defined as the unique relation that enters into such laws (Tooley 1987, 1990a).

The first of these approaches, though deserving of close consideration, offers an account of causation in which no concept of probability plays any role. I shall therefore focus, in what follows, on the second approach.

5.1 Causation and asymmetric probability relations

The basic idea that underlies this approach is that there are certain connections between causation, on the one hand, and prior and posterior probabilities on the other, and the connections in question will emerge if one considers the following case. Let *S* be some very simple type of state of affairs, and *T* a very complex one. (*S* might be a momentary instance of redness, and *T* a state of affairs that is qualitatively identical with the total state of our solar system at the beginning of the present millennium.) In the absence of other evidence, one should surely view events of type *S* as much more likely than events of type *T*. Suppose that one learns, however, that events of type *S* are always accompanied by events of type *T*, and vice versa, and that this two-way connection is nomological. Then one's initial probabilities need to be adjusted, but exactly how this should be done is not clear. Should one assign a lower probability to states of affairs of type S, or a higher probability to states of affairs of type T, or both? And precisely how should the two probabilities be changed?

Contrast this with the case where one learns, instead, that events of type *S* are causally sufficient and causally necessary for events of type *T*. In this case, it is surely clear what one should do: one should adjust the probability that one assigns to events of type *T*, equating it with the probability that one initially assigned to events of type *S*. Conversely, if one learns that events of type *T* are causally sufficient and causally necessary for events of type *S*, then the thing to do is to adjust the probability that one assigns to events of type *S*, equating it with the probability that one initially assigned to events of type *T*.

The relationships between prior probabilities and posterior probabilities are very clear in the case where events of one type are both causally sufficient and causally necessary for events of some other type. But to arrive at the desired postulates, we need to shift, first, to the case where events of one type are causally sufficient, but not causally necessary, for events of some other type, and then we need to generalize to the case where, instead of events of one type being causally sufficient for events of another type, there is only a certain probability that an event of the one type will causally give rise to an event of the other type.

In the case where events of type *S* were both causally sufficient and causally necessary for events of type *T*, the idea was that the posterior probability of an event of type *S*, relative to that causal relationship, was equal to the prior probability of an event of type *S*. When one shifts to the case where an event of type S is causally sufficient for an event of type *T*, the relevant postulate giving the posterior probability of an event of type *S* is as follows:

$$(\mathrm{P}_1)\ \mathrm{Prob}(Sx/\, L(C, S, T)) = \mathrm{Prob}(Sx)$$

where '$L(C, S, T)$' says that it is a law that, for any x, x's being S causes x to be T.

What about the posterior probability of an event of type T? The postulate covering this can be arrived at by starting from the following analytic truth:

$$\begin{aligned}\mathrm{Prob}(Tx\ L(C, S, T)) = {} & \mathrm{Prob}(Tx/Sx\ \&\ L(C, S, T)) \times \mathrm{Prob}(Sx/L(C, S, T)) + \\ & \mathrm{Prob}(Tx/{\sim}Sx\ \&\ L(C, S, T)) \times \mathrm{Prob}({\sim}Sx/\, L(C, S, T))\end{aligned}$$

This then simplifies to:

$$\begin{aligned}\mathrm{Prob}(Tx\ L(C, S, T)) = {} & \mathrm{Prob}(Tx/Sx\ \&\ L(C, S, T)) \times \mathrm{Prob}(Sx) + \mathrm{Prob}(Tx/{\sim}Sx \\ & \&\ L(C, S, T)) \times \mathrm{Prob}({\sim}Sx)\end{aligned}$$

in view of (P_1), plus an immediate corollary of (P_1), namely, $\mathrm{Prob}({\sim}Sx\ L(C, S, T)) = \mathrm{Prob}({\sim}Sx)$.

In addition, it is clearly an analytic truth that $\mathrm{Prob}(Tx/Sx\ \&\ L(C, S, T)) = 1$, so that we can simplify further to:

$$\mathrm{Prob}(Tx/\, L(C, S, T)) = \mathrm{Prob}(Sx) + \mathrm{Prob}(Tx/{\sim}Sx\ \&\ L(C, S, T)) \times \mathrm{Prob}({\sim}Sx)$$

It would seem, however, that if there is no event of type *S*, then the probability of an event of type *T* should not be altered by its being a law that events of type *S* cause events of type *T*. So the following would seem to be a reasonable postulate:

$$(P_2)\ \mathrm{Prob}(Tx/{\sim}Sx\ \&\ L(C, S, T)) = \mathrm{Prob}(Tx/{\sim}Sx)$$

If that is right, one can then move on to the following formula for $\mathrm{Prob}(Tx\, /L(C, S, T))$:

$$(P_3)\ \mathrm{Prob}(Tx/L(C, S, T)) = \mathrm{Prob}(Sx) + \mathrm{Prob}(Tx/{\sim}Sx) \times \mathrm{Prob}({\sim}Sx)$$

The idea, in short, is that in the case of non-probabilistic causal laws, the relations between prior and posterior probabilities are expressed by the two basic principles – (P_1) and (P_2) – along with the derived principle (P_3).

The final step involves generalizing these principles to cover the case of probabilistic causal laws. In the case of (P_1) and (P_2), we need merely replace '$L(C, S, T)$' by '$M(C, S, T, k)$', where the latter says that it is a law that, given an event of type S, the probability that that event causes an event of type T is equal to k. So we have the following two postulates:

(Q_1) $\mathrm{Prob}(Sx/M(C, S, T, k)) = \mathrm{Prob}(Sx)$

(Q_2) $\mathrm{Prob}(Tx/{\sim}Sx \;\&\; M(C, S, T, k)) = \mathrm{Prob}(Tx/{\sim}Sx)$

The principle that is the probabilistic analogue of (P_3) can then be derived as follows. First, given that it is a logical truth that

$$\begin{aligned} M(C, S, T, k) \leftrightarrow\; & Sx \;\&\; C(Sx, Tx) \;\&\; M(C, S, T, k) \text{ or } Sx \;\&\; {\sim}C(Sx, Tx) \;\&\\ & M(C, S, T, k) \text{ or } {\sim}Sx \;\&\; C(Sx, Tx) \;\&\; M(C, S, T, k) \text{ or}\\ & {\sim}Sx \;\&\; {\sim}C(Sx, Tx) \;\&\; M(C, S, T, k) \end{aligned}$$

– where '$C(Sx, Tx)$' says that Sx causes Tx – and given that the disjuncts are all mutually exclusive, it must be an analytic truth that

$$\begin{aligned} \mathrm{Prob}(Tx\ M(C, S, T, k)) = \;& [\mathrm{Prob}(Tx/Sx \;\&\; C(Sx, Tx) \;\&\; M(C, S, T, k)) \times\\ & \mathrm{Prob}(Sx \;\&\; C(Sx, Tx)/M(C, S, T, k))] +\\ & [\mathrm{Prob}(Tx/Sx \;\&\; {\sim}C(Sx, Tx) \;\&\; M(C, S, T, k)) \times\\ & \mathrm{Prob}(Sx \;\&\; {\sim}C(Sx, Tx)/M(C, S, T, k))] +\\ & [\mathrm{Prob}(Tx/{\sim}Sx \;\&\; C(Sx, Tx) \;\&\; M(C, S, T, k)) \times\\ & \mathrm{Prob}({\sim}Sx \;\&\; C(Sx, Tx)/M(C, S, T, k))] +\\ & [\mathrm{Prob}(Tx/{\sim}Sx \;\&\; {\sim}C(Sx, Tx) \;\&\; M(C, S, T, k) \times\\ & \mathrm{Prob}({\sim}Sx \;\&\; {\sim}C(Sx, Tx)/M(C, S, T, k))] \end{aligned}$$

This can then be simplified by making use of (Q_1) and (Q_2) together with the following relationships:

(1) $\mathrm{Prob}(Tx/Sx \;\&\; C(Sx, Tx) \;\&\; M(C, S, T, k)) = 1$ since $C(Sx, Tx) \leftrightarrow Tx$

(2) $\mathrm{Prob}(Sx \;\&\; C(Sx, Tx)/M(C, S, T, k)) = \mathrm{Prob}(C(Sx, Tx)/Sx \;\&\; M(C, S, T, k)) \times \mathrm{Prob}(Sx\ /M(C, S, T, k))$

(3) $\mathrm{Prob}(C(Sx, Tx)/Sx \;\&\; M(C, S, T, k)) = \mathrm{k}$

(4) $\mathrm{Prob}(Sx \;\&\; {\sim}C(Sx, Tx)/M(C, S, T, k)) = \mathrm{Prob}({\sim}C(Sx, Tx)/Sx \;\&\; M(C, S, T, k)) \times \mathrm{Prob}(Sx\ /M(C, S, T, k))$

(5) $\text{Prob}(\sim C(Sx, Tx)/Sx \;\&\; M(C, S, T, k)) = (1 - \text{k})$

(6) $\text{Prob}(\sim Sx \;\&\; C(Sx, Tx)/M(C, S, T, k))] = 0$ since $C(Sx, Tx) \leftrightarrow Sx$

(7) $\text{Prob}(Tx/\sim Sx \;\&\; \sim C(Sx, Tx) \;\&\; M(C, S, T, k) = \text{Prob}(Tx/\sim Sx \;\&\; M(C, S, T, k)$ since $\sim Sx \leftrightarrow \sim Sx \;\&\; \sim C(Sx, Tx)$

(8) $\text{Prob}(\sim Sx \;\&\; \sim C(Sx, Tx)/M(C, S, T, k)) = \text{Prob}(\sim C(Sx, Tx)/\sim Sx \;\&\; M(C, S, T, k)) \times \text{Prob}(\sim Sx/M(C, S, T, k))$

(9) $\text{Prob}(\sim C(Sx, Tx)/\sim Sx \;\&\; M(C, S, T, k)) = 1$

The result of making the relevant substitutions is then as follows:

$$\begin{aligned}\text{Prob}(Tx, M(C, S, T, k)) &= [1 \text{ x k} \times \text{Prob}(Sx)] + [\text{Prob}(Tx/Sx \;\&\; \sim C(Sx, Tx) \;\&\\ &\quad M(C, S, T, k)) \times (1 - \text{k}) \times \text{Prob}(Sx)] +\\ &\quad [\text{Prob}(Tx/\sim Sx \;\&\; M(C, S, T, k) \times 0] +\\ &\quad [\text{Prob}(Tx/\sim Sx \;\&\; \sim C(Sx, Tx) \;\&\; M(C, S, T, k) \times\\ &\quad 1 \times \text{Prob}(\sim Sx)]\\ &= [\text{k} \times \text{Prob}(Sx)] + [\text{Prob}(Tx/Sx \;\&\; \sim C(Sx, Tx) \;\&\\ &\quad M(C, S, T, k)) \times (1 - \text{k}) \times \text{Prob}(Sx)] +\\ &\quad [\text{Prob}(Tx/\sim Sx) \times \text{Prob}(\sim Sx)]\end{aligned}$$

Then, to arrive at the final proposition, we need a principle which says that the probability that *Tx* is the case, given that *Sx* is the case, and that $M(C, S, T, k)$ is the case, but that *Sx* does not cause it to be the case that *Tx*, is just equal to the probability of *Tx* given *Sx* alone. So let us introduce the following postulate:

(Q_3) $\text{Prob}(Tx/\text{Sx} \;\&\; \sim C(Sx, Tx) \;\&\; M(C, S, T, k)) = \text{Prob}(Tx/\sim Sx)$

This allows us to arrive at the following important, derived proposition:

(Q4) $\text{Prob}(Tx, M(C, S, T, k)) = [\text{k} \times \text{Prob}(Sx)] + [(1 - \text{k}) \times \text{Prob}(Tx/Sx) \times \text{Prob}(Sx)] + [\text{Prob}(Tx/\sim Sx) \times \text{Prob}(\sim Sx)]$

The idea is then that postulates (Q_1) through (Q_4) – or simply (Q_1) through (Q_3), given that (Q_4) is derivable from the other three – serve to define implicitly the relation of causation. That implicit definition can then be converted into an explicit one by using one's preferred approach to the definition of theoretical terms. So, for example, if one adopted a Ramsey/Lewis approach, one would first replace the three descriptive terms '*C*', '*S*' and '*T*' by variables ranging over properties and relations. Next, since it is only '*C*' that one wants to define, one affixes two universal quantifiers to the front of the resulting open sentence containing the variables that one put in place of '*S*' and '*T*', so that one has an open sentence with only the one free variable – namely, the one that was used to replace all occurrences of '*C*'. The relation of causation can then be defined as that unique relation between states of affairs that satisfies the open sentence in question.

5.2 The merits of this account

5.2.1 In comparison to an analysis in terms of relative frequencies

The advantages of the analysis of causation just set out emerge very clearly if one considers the problems that confronted the two reductionist approaches. Let us begin, then, with the objections directed against an account in terms of relative frequencies. First of all, a number of problems for the latter account arise from the fact that, according to it, the direction of causation supervenes upon patterns in events – specifically, upon the direction of open forks. Because of this, that account could not handle such possibilities as accidental and non-accidental open forks involving common effects, temporally inverted, twin universes, and simple, temporally symmetric worlds that contain causally related events. By contrast, when causation is viewed as a theoretically defined relation between states of affairs, the different relationships that are set out in (Q_1) and (Q_4) between posterior probabilities and prior probabilities, in the case of causes and effects, ensure that the relation of causation possesses an *intrinsic* direction, rather than a direction supervening upon any patterns in events. Because of this, neither accidental nor non-accidental open forks involving common effects pose any problem. Similarly, nothing precludes there being either temporally inverted, twin universes, or simple, temporally symmetric worlds that contain causally related events.

The same is true in the case of simple, probabilistic, temporally non-symmetric worlds, but here there is the additional advantage that (Q_1) and (Q_4) provide the basis for a justification of one's intuitive judgements about the likely direction of causation in such worlds, since one can show that it is much more likely that the direction of causation runs from the very simple events to the extremely complex ones, rather than in the opposite direction.

Second, there were underdetermination objections, based on situations that are, as far as patterns in events go, causally ambiguous – such as the case where there are excellent theoretical grounds for holding that the probability that an event of type *P* will give rise, directly, to an event of type *Q* is 0.75, where one finds that events of type *P* are directly followed by events of type *Q* with a probability of about 0.76, and where one knows that events of type *Q* occur uncaused with just the frequency that would make it likely, given a law that events of type *P* cause events of type *Q* with a probability of 0.75, that events of type *P* will be immediately followed by events of type *Q* with a probability of about 0.76. If one attempts to analyse causation in terms of relative frequencies, one is forced to the unintuitive conclusion that such cases are logically impossible, since there are no non-causal states of affairs that can distinguish between the cases where an event of type Q has been caused by an event of type P, and the cases where the event of type Q has followed an event of type P, but has not been caused by it. By contrast, if causation is a theoretically defined relation between states of affairs, this possibility poses no problem at all.

Finally, there was the most central and crucial objection of all, namely, that directed against the claim that causes – or, at least direct causes – must raise the probabilities of their effects. Here the problem was that an event of type C might have caused an event of type E, but there might have been another type of event, D, such that, first, the probability of the occurrence of an event of type E is greater, given an event of type D, than given an event of type C, and, second, if the event of type C had not been present, an event of type D would have been.

The view that causation is a theoretically defined relation between states of affairs is, by contrast, perfectly compatible with the idea that, had a certain cause been absent, a more efficacious cause would have been present, since this approach to causation does not entail that a direct cause must raise the probability of its effect in the way claimed by probabilistic, reductionist analyses of causation.

In response, it might be objected that there is surely something intuitively very appealing about the idea that a cause makes its effect more likely. The answer to this, however, is simply that causes do make their effects more likely, but not in the way claimed by reductionist analyses.

The correct account of probability-raising by causes follows very quickly, in fact, from the important principle that we saw was entailed by (Q_1) through (Q_3), namely:

$$(Q_4)\ \text{Prob}(Tx, M(C, S, T, k)) = [\text{k} \times \text{Prob}(Sx)] + [(1 - \text{k}) \times \text{Prob}(Tx/Sx) \times \text{Prob}(Sx)] + [\text{Prob}(Tx/{\sim}Sx) \times \text{Prob}({\sim}Sx)]$$

The derivation is as follows:

First, it is a theorem of probability theory that

$$(1)\ \text{Prob}(Tx) = [\text{Prob}(Tx/Sx) \times \text{Prob}(Sx)] + [\text{Prob}(Tx/{\sim}Sx) \times \text{Prob}({\sim}Sx)]$$

Second, for the 'Tx' that we are considering here, 'Tx' is not logically entailed by 'Sx', nor is it the case that it is logically necessary that Tx. Third, if 'Tx' is not logically entailed by 'Sx', and it is not the case that it is logically necessary that Tx, then the following is true:

$$(2)\ \text{Prob}(Tx/Sx) < 1$$

(Here it is crucial that a distinction is drawn between probabilities that are precisely equal to one, and probabilities that are merely infinitesimally close to one.)

Fourth, if k were precisely equal to zero in '$M(C, S, T, k)$', so that events of type S causally give rise with probability zero to events of type T, then it would not be true that events of type S cause events of type T. So we can assume that

(3) $k > 0$

It then follows from (2) and (3) that:

(4) $[\mathrm{k} \times \mathrm{Prob}(Sx)] + [(1 - \mathrm{k}) \times \mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] > [\mathrm{k} \times \mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] + [(1 - \mathrm{k}) \times \mathrm{Prob}(Tx/Sx) \text{ x } \mathrm{Prob}(Sx)]$

It is also true, however, that:

(5) $[\mathrm{k} \times \mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] + [(1 - \mathrm{k}) \times \mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] = [\mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)]$

Statements (4) and (5) together then give us:

(6) $[\mathrm{k} \times \mathrm{Prob}(Sx)] + [(1 - \mathrm{k}) \times \mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] > [\mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)]$

This, in turn, together with (Q_4), yields:

(7) $\mathrm{Prob}(Tx, M(C, S, T, k)) > [\mathrm{Prob}(Tx/Sx) \times \mathrm{Prob}(Sx)] + [\mathrm{Prob}(Tx/{\sim}Sx) \times \mathrm{Prob}({\sim}Sx)]$

Finally, (7), together with (1), gives us the following result:

(Q_5) $\mathrm{Prob}(Tx, M(C, S, T, k)) > \mathrm{Prob}(Tx)$, provided $k > 0$

This says that causes do raise the probabilities of their effects, in the following way: the probability of Tx, given only that it is a law that events of type S give rise, with some non-zero probability k to events of type T, is greater than the a priori probability of Tx.

My basic thesis concerning the raising of the probabilities of effects by their causes is, accordingly, that this is the case only in the sense stated by principle (Q_5).

5.2.2 In comparison to an analysis in terms of objective chances

Next, let us consider how the approach according to which causation is a theoretically defined relation between states of affairs compares with an analysis of causation in terms of objective chances. To begin with, then, we saw that the idea that objective chances were ontologically ultimate was exposed to at least four serious objections. First, that idea entails that there can be relations of logical entailment between temporally distinct, intrinsic states of affairs. By contrast, when one defines the relation of causation in the manner indicated above, one can then go on, first, to define causal laws as laws that involve the relation of causation, and then, second, to define objective chances in terms of causal laws plus non-causal properties and relations. When this is done, the existence of objective chances does not entail any logical relations between temporally distinct states of affairs.

A second and related objection was that when causation is analysed in terms of objective chances, it turns out that rather than laws connecting states of affairs existing at different times, what one has are laws connecting states of affairs at one and the same time, plus logical connections between temporally distinct states of affairs. This means that one is confronted with the puzzle of why, if there can be laws of nature connecting simultaneous states of affairs, it should be impossible for there to be laws connecting states of affairs that exist at different times.

When causation is analysed as a theoretically defined relation between states of affairs, this problem does not arise: both basic causal laws linking states of affairs at different times, and basic laws of co-existence linking states of affairs at a single time, are logically possible.

A third objection was that a state of affairs at a single instant may involve a non-denumerable infinity of objective chances. If objective chances are ontologically ultimate, that means that the momentary state of affairs involves an infinite number of distinct, intrinsic properties. But if objective chances, instead, supervene on non-causal properties and relations plus causal laws, then there is no need for any infinity of properties. Indeed, an infinity of objective chances may even supervene upon a single intrinsic property plus a single causal law.

A fourth objection to ontologically ultimate, objective chances was that there are objective chances that are, intuitively, perfectly compatible, but that would be incompatible if objective chances were ontologically ultimate. We also saw that the objective chances in question are perfectly compatible when one analyses objective chances in terms of causal laws plus non-causal properties and relations.

Next, there were two objections that arose in the case of a relative frequencies approach to causation that also apply to an analysis of causation in terms of objective chances. First, the latter approach is also exposed to underdetermination objections, since the arguments that show that causal relations between states of affairs do not supervene upon causal laws plus non-causal properties and relations also show that the situation is not changed if one adds objective chances to the proposed supervenience base. Second, an analysis of causation in terms of objective chances incorporates the requirement that at least a direct cause of an effect must raise the objective probability of its effect, and so this approach is also exposed to the objection that there are situations where if a certain cause had been absent, a more efficacious cause would have been present.

By contrast, as we have just seen, neither of these two objections poses any problem at all for the view that causation is a theoretically defined relation between states of affairs. On the contrary, as regards the second of these objections, it is one of the great strengths of the latter account that it can provide a correct account of the one and only way in which causes do raise the probabilities of their effects.

Finally, there was the objection that an analysis of causation in terms of objective chances *immediately* rules out a causal analysis of the direction of time, since, in general, objective chances of an event of a given type, E, are not chances that the event will occur at some time or other, nor even chances that an event of type E will occur at a certain temporal distance: they are, instead, chances that an event of type E will occur at a certain temporal distance *and* in a certain temporal *direction*. Ontologically ultimate, objective chances presuppose, therefore, the relation of temporal priority, and so if one analyses causation in terms of such objective chances, a causal analysis of the direction of time is ruled out. By contrast, when causation is viewed as a relation between states of affairs that is to be defined via the theory set out above, the idea of a causal account of the direction of time remains an open possibility.

6 Summing up

In the preceding sections, I have argued against two accounts of the relation between probability and causation, and in favour of another. In particular, I have attempted, first of all, to establish the following claims:

(1) Reductive analyses of causation in terms of relative frequencies are untenable.
(2) Reductive analyses of causation in terms of ontologically ultimate, objective chances are also open to decisive objections.

Second, I have also tried to make plausible the following two claims:

(3) There are necessary connections between causation and logical probability, and those connections are captured by postulates (Q_1) through (Q_3), and, in a more explicit fashion, by the two derived propositions (Q_4) and (Q_5).
(4) The correct analysis of the relation of causation is given by a theoretical-term style definition based upon the theory that consists of postulates (Q_1) through (Q_3) – supplemented, if one prefers, by (Q_4) and/or (Q_5).

Notes

1 See, for example, Dretske 1977; Tooley 1977, 1987: section 2.1.1; Armstrong 1983, esp. ch. 1–5; and Carroll 1994.
2 For three other ‘underdetermination arguments, see Tooley (1990b).
3 For a precise formulation of the last condition, see Mellor (1995: 38–43).
4 Mellor’s own formulation (1995: 175–9) is different, and considerably more complicated.

Chapter 7

Analysing chancy causation without appeal to chance-raising

Stephen Barker

Introduction

Must counterfactual analyses of causation appeal to chances and chance-raising in order to tame indeterministic causation? It is generally thought so.[1] Against the grain, I contend that appeal to chance-raising is not required to analyse chancy causation. In Section 1 below I argue that the standard cases motivating the chance-raising analysis – cases such as bombardments of radioactive atoms causing the decay of those atoms – should be treated as instances of preemption. Such cases, I urge, are open to exactly the same kind of analysis as cases of preemption in a deterministic setting. With that thought in mind, I set out to provide a unified counterfactual analysis of causation for both deterministic and indeterministic cases, without appeal to chance-raising. In Section 2, I outline an initial sketch for the deterministic case of the theory. The account is a descendant of Lewis's quasi-dependence analysis but expressed in terms of counterfactual embedding. I show how this theory copes with preemption, trumping and over-determination. In Section 3, I broach the vexed issue of the non-transitivity of cause. I accept that cause is not transitive, and modify the theory given in Section 2 to derive the final theory. The overall picture is roughly this: c caused e if and only if c and e occur and had c not occurred a disposition to a causal path issuing in e would not have been manifested in the circumstances. I show in Section 4 how this theory applies to chancy causation.

1 Chance-raising and processes

Say an unstable atom, *b*, has a residual chance of decay in the next instant determined by its half-life. Suppose *b* is bombarded by a photon, *p*, and this raises the chance of *b*'s decay. As a matter of fact, *b* decays in the next instant, emitting energy as it does so. We might judge that the bombardment caused *b* to decay. However, it is not the case that if *p* had not hit *b*, *b* would not have decayed, since *b* had a residual chance of decaying, and as such might have decayed even if it hadn't been bombarded. The simple counterfactual dependency of effect on cause does not obtain here, but neither, it seems, does any other counterfactual dependency condition, say, involving dependency with respect to other events. How then do we explain the causal judgement that *p*'s bombarding *b* caused *b* to decay? Lewis (1986: 175–84) proposes that the causation is constituted by chance-raising. He defends the thesis CR:

> CR: For all occurring events *c* and *e*, *c* caused *e* if [a]–[c] hold:
> [a] The chance at the time of *c*, t_c, of *e* is ω.
> [b] Had *c* not occurred, the chance at t_c of *e* would have been θ.
> [c] ω is significantly larger than θ.

In the case under scrutiny, *p*'s bombarding *b* raised the chance of decay; had there been no bombardment, the chance of *b*'s decay would have been significantly lower. Given CR, we explain our judgement that bombardment caused decay.

Is it plausible to think that particular cases of causation are constituted by chance-raising – as specified in CR? There is a good reason to think not. Till we know more about the case, we can only assign a certain degree of likelihood to the proposition that *p*'s hitting *b* caused *b*'s decay. I illustrate this thought with some fairy-tale physics. Say that *b* has two separate energy systems. One is *b*'s residual energy state, responsible for its residual chance of decay. The other is the system that is activated by bombardment, which, in an activated state, also contributes to the chance of decay. These systems are causally autonomous from each other. Thus one system can be activated by photon bombardment but it is the residual system that is responsible for the decay. Schematically, this possibility is represented by Figure 7.1.

The lower dark circle is the residual system. The upper dark circle is the system activated by bombardments; the circle containing both is *b*. The lighter shaded arrow indicates the upper system is inactive in triggering decay, the darker arrow is the path of causation from the residual state. If Figure 7.1 represents the structure of the bombardment situation, then from the fact that *p* bombarded *b*, *b* decayed, and bombardment raised the chance of decay, it does not follow that the bombardment caused decay, despite the fact that there is chance-raising.[2]

To argue that the bombardment did cause the decay, we need to assume that *p* itself altered *b*'s energy state, ensuring that its residual state was no longer extant. We need a structure such as that shown in Figure 7.2.

What Figure 7.2 represents is a kind of *non-gappy* causal process leading from cause to effect. Let us leave non-gappy causal processes on an intuitive level at this

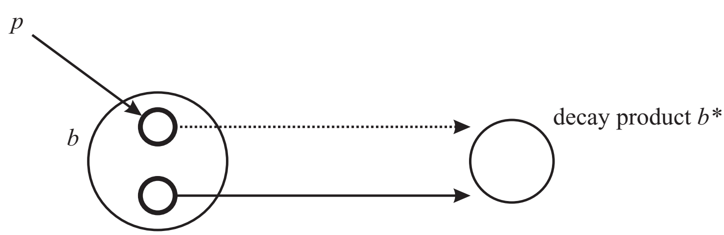

Figure 7.1

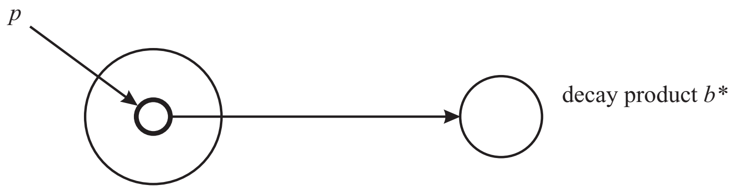

Figure 7.2

stage.[3] In Figure 7.1 there is no non-gappy process from cause to effect; the two sub-systems are unconnected. Unless we can be assured that a process of the kind exhibited in Figure 7.2 is present in the bombardment case, we can only say that it was likely to some degree that the bombardment caused the decay.[4]

These conclusions are supported by many other counterexamples to CR. Take Ramachandran's example (Ramachandran, forthcoming). Say that there are four neurons as shown in Figure 7.3.

A and *B* fire, but *C*, very improbably, does not, but *D* spontaneously fires. In this case *A*'s firing does not cause *D*'s firing, but counterfactually it raises its chance. So CR is wrong.

In conceiving non-gappy processes, we don't always have to think of chances, such as chances of decay, as being underpinned by categorical properties – such as being-in-energy-state-*E*. We can define processes just in terms of primitive facts of chance, which do not hold in virtue of any categorical fact.[5] If we are thinking of chances as single-case objective chances, then they are properties of objects – call these properties *propensities* – which may or may not be manifested. Say we were to think of the bombardment case in these terms. Then we conceive of bombardment as bringing about the instantiation of a propensity in *b*. We can think of the process in this case as having a structure identical to that in Figure 7.2, except that instead of *b*'s being in an energy state we have *b*'s possessing a propensity. We can also think of a process in terms of propensities with the structure of Figure 7.1. In this case, there are propensities that are instantiated by parts of *b*; the separate energy systems are replaced by components of *b* that have distinct propensities. Thus, even if we were to reject the idea that there were energy states underpinning chances, accepting instead that there are just propensities present, to judge that there is causation afoot we still need more than chance-raising; we need, at the very least, a non-gappy causal process defined in terms of events and objects instantiating propensities.[6]

The importance of processes suggests that we ought to replace CR with a sufficient condition for causation along the following lines: *c* caused *e* if *c* and *e* occurred, *c* raises the chance of *e*, and *c* is linked to *e* by a causal process.[7] But, in

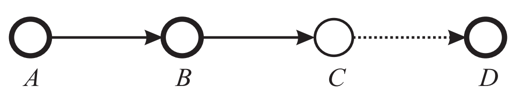

Figure 7.3

fact, reference to chance-raising is now entirely otiose; it does no work that a straight counterfactual dependency cannot do.[8] The argument for this proceeds as follows.

The judgement that *p*'s bombarding *b* caused *b*'s decay depends upon postulation of a process comprised by energy states, or propensities. If an atom's having an energy state, or propensity, can be a component of a causal process, it is also a potential cause.[9] Consider now the case where *b* has not been bombarded by a photon, and has an energy state *E*1. If *b* decays, then *b*'s being in *E*1 is part of the causal process leading to decay. This is so for the following reason. If a causal process links one event, e_1, with another, e_2, then, either e_1 caused e_2 or e_2 caused e_1. In the simple decay case, the decay did not cause the energy state, so *b*'s being in *E*1 must have caused the decay. How does the modified chance-raising analysis explain the fact that *b*'s being in *E*1 caused *b*'s decay? It must be that *b*'s being in *E*1 raised the chance of the decay event and was linked to it by a process. That is, (1) holds:

(1) [a] The actual chance of decay was ω.
[b] If *b* had not been in *E*1, the chance of decay would have been θ.
[c] ω is significantly greater than θ.
[d] A process links *b*'s being in *E*1 and the decay event.

But this won't work. If we take the antecedent of (1)[b] to be tantamount to supposing that *b* is in some other energy state than *E*1, the value ω will be higher than θ, since all the other energy states will be ones with greater propensities for decay. In short, we cannot explain the causation by appeal to chance-raising and process, or so it seems.[10]

Perhaps, there is a way out for the chance-raising account. The problem is with the counterfactual (1)[b]. The relevant supposition cannot be that *b* was not in *E*1. Rather, the content of the supposition needs to be something like this: suppose that *b* was not in *E*1 or any other energy state. With *b* in no energy state at all it has no propensity to decay. Thus, that *b*'s being in *E*1 caused the decay might be constituted by these conditions:

(2) [a] The actual chance of decay was ω.
[b] If *b* had not been in *E*1, or any other state, the chance of decay would have been θ.
[c] ω is significantly greater than θ.
[d] A process links *b*'s being in *E*1 and the decay event.

Condition (2)[c] holds since ω will be zero. Of course, the counterfactual (2)[b] is a counterlegal in which we conceive of *b*'s being in the physically impossible situation of lacking any energy state. (2)[b] is true because decay requires a prior energy state of some kind, and so the absence of any energy state implies that there is no decay. We might wonder how this kind of suppositional exercise – supposing that *b* is not in any other energy state – arises from a counterfactual analysis of causation. Let us leave this question aside for the moment (I take it up again in Section 4 on p. 132).

If we accept this defence of the chance-raising analysis, we have shown it can explain our causal judgements in these simple non-bombardment cases, but only at the price of making reference to chance-raising redundant. In (2), the chance θ is zero, because there is no decay event possible under the supposition that *b* is in no energy state at all. The would-counterfactual, 'If *b* had not been in *E*1 (or any other state), then *b* would not have decayed', is also true. So, we could have got the same explanatory effect just by appeal to this would-counterfactual. There is no reason to invoke chance-raising.[11]

If we can deal with the simple decay case in this fashion, what about the bombardment case? We can analyse the bombardment case as an instance of preemption. The actual cause of *b*'s decay is *p*'s hitting *b*, but the preempted back-up cause is *b*'s being in its residual energy state. Whatever story we tell in general about preemption – for the deterministic case – can be told here. If we can tell that story, the result is a unified counterfactual analysis of both deterministic and indeterministic causation.

2 The ED analysis: first try

I intend, then, to proceed with the conclusion in hand that we can, in principle, treat indeterministic cases in terms of straight counterfactuals, if we appeal to energy states – or propensities – and if we have the right account of preemption. What general kind of counterfactual analysis of causation should be adopted that can explain the causal preemption in the bombardment example? In what follows I offer an account that not only deals with preemption, but also, I argue, with trumping, distinguishing hasteners and delayers from causes, and over-determination. I modify this account in Section 3 to deal with the non-transitivity of cause. In Section 4, I return to indeterminism, and flesh out, in the form of a theory, the sketch given at the end of the last section.

How do we deal with preemption? The proposal I am putting forward stems from the idea that causation is comprised by a certain kind of event-counterfactual dependency, but one that may fail to be revealed due to interfering external factors. Preemption is an example of this. Consider the following classic case. Fred goes into the desert. He has two enemies, Zack and Mack: Zack aims to poison him by putting poison in his water-bottle; Mack aims to cause him to die of dehydration by punching a hole in his bottle. Both these events occur. The water runs out of Fred's bottle and Fred dies of dehydration. Thus, Mack's punching a hole in Fred's bottle caused his death, even though there is no simple counterfactual dependency of death on Mack's holing the bottle. Structurally the case can be shown by Figure 7.4.

Nevertheless, we recognize that there is a process of the right kind from the holing to death. By isolating this process from its actual context we can make that dependency manifest. A simple way of doing this is to suppose that the preempted cause had not occurred. If there had been no poisoning, then the holing-death process would still

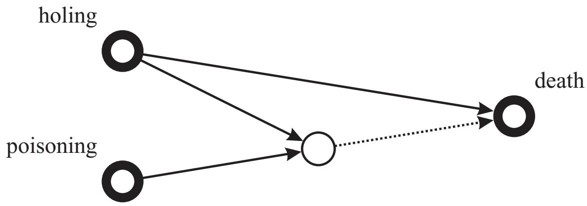

Figure 7.4

have obtained. Moreover, the dependency – had there been no holing, there would have been no death – would have held under those conditions. Thus, one might argue, holing was the cause of death for this reason. But can't one argue equally in the opposite direction that had there been no holing, then there would still have been a poisoning and a death, and moreover a dependency of the latter on the former? The answer is no; there is no completed process in this case. If there had been no holing, then the poisoning-death process would have been completed, that is, events that actually did not occur would have to have occurred to get the effect; had the holing and these events not occurred, there would have been no death. That is the asymmetry, and why poisoning is not a cause whereas holing is. We can sum up the theory thus:

The Embedded Dependency – ED – account: (c caused e) iff
[1] O(c) and O(e).
[2] There are events, conditions or states, f – possibly non-obtaining such that:
a. $(\neg O(f) > (O(c) \& O(e)))$ b. $(\neg O(f) > (\neg O(c) > \neg O(e)))$.
[3] No non-obtaining event, condition or state, g, is such that:
a. $(\neg O(f) > O(g))$ b. $((\neg O(f) \& \neg O(g)) > \neg O(e))$[12].

The idea is that the causation is ultimately based on counterfactual dependency. That is given in [2b]. For the dependency to reveal itself we cut out the alternative paths leading to the effect. But the paths must be complete in the actual world, a fact guaranteed by [3]. We apply the ED account to the poisoning case as follows – here poisoning, holing and death are the three events concerned. It is the case that (holing caused death) since:

[1] O(holing) and O(death).
[2] There is an event, poisoning, such that:
a. (¬O(poisoning) > (O(holing) & O(death)))
b. (¬O(poisoning) > (¬O(holing) > ¬O(death))).
[3] No non-occurring event, condition or state, g, is such that:
a. (¬O(poisoning) > O(g)) b. ((¬O(poisoning) & ¬O(g)) > ¬O(death)).

In contrast it is not the case that (poisoning caused death) since although [1] and [2] hold, [3] does not:

[1] O(poisoning) and O(death).
[2] There is an event, holing, such that:
a. (¬O(holing) > (O(poisoning) & O(death)))
b. (¬O(holing) > (¬O(poisoning) > ¬O(death))).

But there is some event, *g*, say, poison enters George such that:

¬O(poison enters George)
((¬O(holing) & ¬O(poison enters George)) > ¬O(death))

The idea here is that the chain from *poisoning* to *death* is incomplete, and a necessary condition for causation fails to obtain. Completion would have come about only if there had been no holing of George's bottle.

That is the ED theory in essence.[13] ED involves no step-wise chains of dependence; it requires dependency of effect on cause, under the hypothetical conditions.

Late preemption

We have looked at early preemption above. Take a case of late preemption, as in Figure 7.5.

In this case, black circles are fired neurons – *B**, *A** and so on mean that *B*, *A*, and so on have fired – and light grey circles unfired neurons, with arrows indicating stimulatory connections, and reverse arrows inhibitors. In this case, the *A**–*E** process occurs, but is faster than the *B**–*E** process. If *E** occurs before *C**, the former inhibits the latter. *A** causes *E**, but if *A** had not occurred, it is not the case that *E** would not have occurred, since *E** would have occurred later. Lewis's step-wise dependency approach, which can cope with early preemption, fails here: *D** depends on *A** but *E** does not depend on *D**.

The ED account applies straightforwardly. That is (*A** causes *E**) because:

[1] O(*A**) and O(*E**).
[2] There is an event, *B**, such that:
a. (¬O(*B**) > (O(*A**) & O(*E**)))
b. (¬O(*B**) > (¬O(*A**) > ¬O(*E**))).
[3] No non-occurring event, condition or state, *g*, is such that:
a. (¬O(*B**) > O(*g*)) b. ((¬O(*B**) & ¬O(*g*)) > ¬O(*E**)).

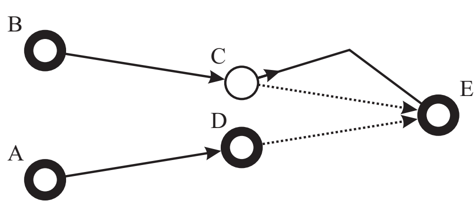

Figure 7.5

In terms of Figure 7.5, the conditions required by ED amount to isolating the A^*–E^* chain by supposing B^* has not occurred, and noting that it is complete. In contrast, if we isolate the B^*–E^* chain by supposing A^* does not occur, we find that it is incomplete. The event C^*, which is non-occurring, must occur for B^* to cause E^*.

Frustration: preempted causes with completed causal paths?

It might be objected that this ED account is doomed since it cannot deal with preemption in which the causal process of the preempted cause is complete. Are there such cases? Noordhof (1998a) argues there are. His example is shown in Figure 7.6.

A fires at time 0 and *B* fires at 1. It takes 2 units of time for the impulse from *B* to reach *D*, and 4 units for *A*'s impulse to reach *D* via *C*. Intuitively the cause of *D*'s firing is *B*'s firing. Noordhof argues that both A^*–D^* and B^*–D^* paths are complete. If so, for an account like ED, there is no way of displaying the asymmetry between A^* and B^*.

My reply is this. Does *D* fire twice? Assume it does not, then, given determinism, there must have been something incomplete in the A^*–D^* path since, if not, *D* would have been necessitated to fire. If so, there is no problem for ED; the path is incomplete. But what of the case where *D* fires twice? Then, *B*'s firing causes the first *D*-firing event, and *A*'s firing the second. One can then argue that there is straightforward counterfactual dependency. If *B* had not fired then the first *D*-firing event would not have occurred.

Over-determination

The ED theory is consistent with effects being over-determined by effects. Say that a revolutionary is to be executed by firing squad. There are four members of the execution party, all with loaded guns. They all aim at the same moment and fire, each killing the revolutionary. Structurally, the situation can be shown by Figure 7.7.

In this case, each *F*-event is a contributory or partial cause. We don't want to say that there was no cause, or that the only cause was disjunctive, since disjunctive events are problematic. We likewise do not want to treat the fusion of the firing of

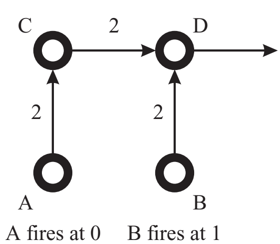

Figure 7.6

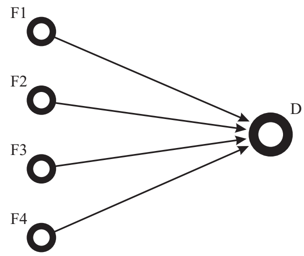

Figure 7.7

$F1$ to $F4$ as the cause, because – contra Lewis, (1986: 212) – there will be no dependency, since removing this event can be done simply by removing one of its parts, say the firing of $F1$, in which case D still occurs.[14] The ED theory explains how each of the F-events is a contributory cause. Thus, suppose that $F2$ to $F4$ had not fired. Then $F1$ would still have fired and there would still have been a death. Moreover, there would have been a counterfactual dependency of D on $F1$ firing. Furthermore, the process-completeness condition [3] is met as well. The ED theory gets the right result.

Hasteners and delayers

Are hasteners causes? Not necessarily. I partially open a window and as a result a ball heading towards the window frame smashes the window earlier than it would have otherwise. Opening the window was a hastener of the smashing, but not a cause of the smashing.[15] ED respects that fact. Under the circumstances, there is no event or condition such that, had it not been the case, there would have been a dependency of the smashing on the window's being opened. According to ED, then, not all hasteners are causes. On the other hand, delayers can be causes. Take our first case of preemption, the poisoning/holing example, with the structure in Figure 7.4. Suppose that holing the water bottle is part of a causal process, death by thirst, which takes much longer than poisoning. So holing the water bottle delayed George's death. Nevertheless, it was a cause of death.

Trumping

Schaffer (2000a) introduced the concept of trumping causes. Say that Merlin casts a spell at 10 a.m. to turn the prince into a frog at midnight. Morgana casts a spell at 11 a.m. to turn the prince into a frog at midnight. There is a law governing spells:

> Spell law: If x casts a spell of the form 'person z will F at time t', on a day y, then if no one else casts a spell earlier on y of the form 'z will G at time t', then x's spell-outcome occurs; z will F at time t.

If this law holds, then Merlin's spell caused the frogification, not Morgana's spell. Trumping cases are puzzling since there is no counterfactual dependency of effect on cause, and the usual methods for dealing with preemption fail.[16] Lewis's (2000) influence theory is meant to account for such cases, but it is not obvious that it does. On the influence account, a sufficient condition for *c* causing *e* is that there is an influence of when, how or whether *e* occurs on when, how or whether *c* occurs. There is no 'whether' dependency of frogification on Merlin-casts, but there is an influence of time and manner. I note, however, that there is an influence of time, though not of manner, of frogification on Morgana-casts. Why is this latter influence not sufficient for causation?

The ED theory accounts for the causation. (Merlin-casts causes frogification) since:

[1] O(Merlin-casts) and O(frogification).
[2] There is an event, Morgana-casts, such that:
(¬O(Morgana-casts) > (O(Merlin-casts) & O(frogification)))
(¬O(Morgana-casts) > (¬O(Merlin-casts) > ¬O(frogification)))
There are no non-occurring *g* such that:
a. (¬O(Morgana-casts) > O(*g*))
b. ((¬O(Morgana-casts) & ¬O(*g*)) > ¬O(frogification)).

So by ED we have causation. On the other hand, it is not the case that (Morgana-casts causes frogification) since, although the first two conditions are met:

[1] O(Morgana-casts) and O(frogification).
[2] There is an event, Merlin-casts, such that:
(¬O(Merlin-casts) > (O(Morgana-casts) & O(frogification)))
(¬O(Merlin-casts) > (¬O(Morgana-casts) > ¬O(frogification))).

The third condition is not met. There is a non-occurring *g* = (Morgana's spell is the first cast on the day) or Morgana-first, for short, such that:

(¬O(Merlin-casts) > (O(Morgana-first)))
((¬O(Merlin-casts) & ¬O(Morgana-first)) > ¬O(frogification))[17]

The particular condition, *g*, here, Morgana-first, is not an event in the standard sense, but it is a fact or state of affairs. So Condition [3] of ED is not met.

3 Are causal paths sufficient for causation?

Is ED a completely satisfactory account of deterministic cause? It isn't. The reason is the non-transitivity of cause. There are a host of examples that illustrate the breakdown of transitivity. Let a causal path between *c* and *e* be a process linking *c* to *e*, where each stage causes the next.[18] If cause were transitive then causal paths would be sufficient for causation. But they are not. Consider, for example, the bomb case – from Yablo (2002) and Hall (2000), credited to Hartry Field: Jane is a healthy woman going about her business and due to go to the doctor the next day for a check-up. A bomb is planted under her work desk. She notices the bomb and leaves the area before it explodes and maims her. She gets a glowing health report from her doctor the next day. There is a causal path from the presence of the bomb to her health the next day, but the presence of the bomb did not cause the next day's health.[19] In Hall's (2000) train-switching case, a train is heading towards the terminus, but between it and the latter is a switching point with tracks A and B. After they diverge they converge at a lower point on the track. Billy is switching between A and B randomly during the day. When the train gets to the switch point he has switched to A. The train goes up the A-track and then, after converging with the main track, goes to the terminus. Billy's switching caused the train to go up the A-track, its going up that A-track caused its later movement along that track, this in turn caused its movement further down the track, and so on. There is then a causal chain linking Bill's switching to the train arriving at the terminus. But intuitively, Billy's switching did not cause the train's arrival at the terminus.[20]

ED validates transitivity. If *c* and *e* are events satisfying ED, then there is a causal path, from *c* to *e* – write this as ($c \rightarrow e$) – and ED takes that as sufficient for *c* to cause *e*.[21] But, as we have seen, the existence of ($c \rightarrow e$) does not imply that *c* caused *e*. The central assumption underpinning ED is wrong. Does this mean that ED has been a complete dead end? Not quite. A causal path ($c \rightarrow e$) though not sufficient for *c*'s causing *e* is necessary. We can build on this fact to provide sufficient conditions of causation by appealing to the notion of a disposition to a causal path.

Where *c* caused *e* there is a disposition in the circumstances, *D*, whose manifestation is a causal path, from the time of *c* to *e*. *D* is a disposition to produce a causal path to *e* if conditions are right. *D*'s manifestations are particular casual paths of a certain kind. If the presence of oxygen caused a match to light, then, *D* is a disposition to causal-paths of the form 'the match lights through striking and combustion'. If Jane's getting out the way of the bomb is a cause of her health the next day, then *D* is a disposition to causal-paths of the form 'being healthy the next day by maintaining a certain level of bodily well-being and intactness'. ED assumes that a cause is simply a necessary component of an actual causal path leading to *e*. But this is wrong. A cause is rather a condition of a disposition *D* being manifested at all. Roughly, *c* causes *e* if and only if had *c* not occurred then the disposition *D*, which is manifested in the path ($c \rightarrow e$), would not have manifested at all. Thus, the presence of oxygen was a cause of lighting because had it not been present, a disposition to there being a causal-path of the form 'lighting-by-combustion' would not have manifested itself. Jane's getting out of the way of the bomb is a cause of her health the next day, because had she not left the area, a disposition of the form 'health-tomorrow-by-health-maintenance' would not have been manifested.

On this conception of causation, cause is not transitive. Occurrence of event *c* may be part of the actual causal path leading to *e*, which is the manifestation of *D*, and so *c* is linked to *e* by a series of intermediate causes. But if *c* had failed to occur, then *D* might still have manifested itself in another path (non-*c*→*e*), where both (non-*c*→*e*) and (*c*→*e*) are manifestations of *D*. So despite the causal path (*c*→*e*), *c* is not a cause of *e*. In the bomb case, the disposition is to causal paths of the form 'Jane is healthy the next day through her maintaining a state of sufficient bodily well-being from an earlier state of health'. Had there been no bomb, there might have been a causal path leading to Jane getting a good health report, which was a path of health-maintenance. So although the bomb's presence is a component of the actual path, it is not required for the manifestation of the path-disposition, and so is not a cause. In the train-case, the disposition is to causal paths of the form 'the train arrives at the terminus by self-propulsion along a track'. If Billy had not switched, then there would have been another path of the same kind – that is, a causal path of the train arriving at the terminus by self-propulsion along a track. In short, in these cases, *c* to non-*c* is mere path-switching.

In summary, where *c* caused *e*, there is a dispositional state to arrive at *c* via some means *x*, where *x* is the type of path; abbreviate this as *D*[*e*-by-*x*]. The new theory of caused, ED+ is:

> ED+: (*c* caused *e*) iff there is a disposition *D*[*e*-by-*x*] that has an actual path (*c*→*e*) as a manifestation such that with respect to this disposition *c* is not a mere path-switcher; had c not occurred, then *D*[*e*-by-*x*] would not have manifested itself.

ED+ depends crucially on the notion of a path-type; the path to *e* by *x*. The laws in operation determine what *x* is. For example, in the bomb case, the laws that govern the process are health-preservation laws of the form:[22]

> L1: If *z* is healthy at *t* and there are no outstanding threats then *z* is healthy at a time after *t*.
>
> L2: If *z* is in a situation that *x* recognizes threatens her bodily well-being, then (ceteris paribus) *z* undertakes certain actions that might issue in threat removal.

In the bomb case a bomb is present, which Jane perceives, and, due to L2, Jane undertakes avoidance behaviour; she leaves the room. Of course, if we treat getting a good health report as the effect, then the causal path would be governed first by laws like L1 and L2, then by another law of the form:

> L3: If *z* is healthy at *t* and *z* is examined by a competent doctor then *z* gets a good health report at a time after *t*.

The process then is governed by two phases corresponding to the laws: the first phase corresponding to the health-preservation laws and the second to the examination-report law.

It is useful to compare cases of transitivity failure with cases of preemption. Take the desert-traveller case discussed in Section 2 on p. 124. The actual causal process that passed the ED test is the (holing→death) path. This path is governed by laws concerning water, gravity, dehydration, and so on. But the holing path is a potential manifestation of a different path-disposition from the poisoning-path. So, holing is not a mere path-switcher, it is a path-maker; if there had been no holing there would not have been any manifestation of *D*[death-by-dehydration]. It might appear, however, that ED+ is refuted by the following case. There are two assassins both using poison to kill the king, one waiting to back up the other. Assassin 1 puts in the poison and the king dies. But if he had not, assassin 2 would have done so. Assassin 1's act caused the death of the king; even though had he failed to put the poison in the king would still have been poisoned.[23] But in this case, don't we simply have a case of path-switching? We have the same poison process, but undertaken by distinct individuals. The reply is that we have a different process. There are two dispositions present, at t_c, the time of the first poisoning. One is a pure poisoning path, to the death of the king, the other is a two-phase path: first phase – the initiation of the back-up poisoner – and second phase a path of poisoning. Because this path has two phases, it is overall a different causal path.

How does the ED+ account cope with over-determination? In the over-determination case discussed on p. 127, there are four distinct causal-paths leading to the death of the revolutionary. These are four distinct manifestations of four distinct path-dispositions – though they have a higher type-identity. Had *F*1 not fired, then the corresponding disposition to a causal path leading to death would not have been manifested. *F*1's firing is, then, a cause of the death, and likewise for each other *F*-event.[24]

3.1 Paths, preventions and omissions

ED+ does not distinguish in any way between causation by omission or prevention and other kinds of causation. ED+'s notion of a causal path is one which allows nodes in the path to be negative states of affairs or omissions. Modifying an example from Schaffer (2001a), say Pam is a saboteur who explodes a bomb, causing a control tower to be destroyed thereby preventing a plane from being warned about a mountain ahead of it. It crashes. Pam's exploding the bomb caused the plane to crash. According to ED+ there is a path here. Roughly, it has the structure: (explosion → control tower destruction → the absence of a warning → the plane's destruction). These events/states of affairs are linked by law-based generalities. ED certifies that the whole path is as a causal path. The whole path is a token of a certain path-type, which would not have been manifested if the explosion had not occurred. Thus, the explosion caused the plane's destruction.

4 Indeterminism again

In Sections 2 and 3, I have described a counterfactual analysis, ED+, that apparently deals with problems of preemption, over-determination, trumping and causal non-transitivity. Let us now return to indeterminism. I have promised a unified treatment of causation: one in which chance-raising plays no role. What we want is to treat cases of indeterministic causation, like the bombardment case, as instances of preemption. I now show how to do this by adopting the ED+ analysis. Take first the bombardment example with which we began. In this case, *p*'s bombarding *b* causes it to decay because the following conditions hold. Roughly, *b* is bombarded by *p*; this caused *b*'s residual energy state, which is *E*1, to change to *E*2, as a consequence of which the chance of decay changes. Given this process, bombardment causes decay. ED+ accounts for this judgement. Here the relevant events are: bombard – the event of *p*'s bombarding *b* – being-in-*E*1 – the event of *b*'s being in its residual energy state *E*1 – and decay – *b*'s decaying. The required necessary condition for the claim (bombard causes decay), that there should be a causal path, holds. That is, the ED condition holds since:

[1] O(bombard) and O(decay).
[2] There are events *f* – in this case, being-in-*E*1–3-N[25] – such that:
  a. (¬O(being-in-*E*1–3-N) > (O(bombard) & O(decay)))
  b. (¬O(being-in-*E*1–3-N) > (¬O(bombard) > ¬O(decay))).
[3] No non-obtaining *g* is such that:
  (¬O(being-in-*E*1–3-N) > O(*g*)) and:
  ((¬O(being-in-*E*1–3-N) & ¬O(*g*)) > ¬O(decay)).

In short, the causation operates on the energy states – the grounds for the chances concerned – whereby the bombardment is responsible for an energy state, *E*2, which renders decay likely. It raises the chance of decay, but does not cause it in virtue of that fact. Rather, it is causal because decay depends upon *b*'s possessing *E*2, in the appropriate way.

But do we have causation according to ED+? For that to hold, it must be that the path so defined is the manifestation of a path-disposition *D*[decay-by-*x*], where *x* is the path-type and bombardment is not a mere path-switcher. The *x*-type is fixed by the probabilistic law that governs the actual process: a law concerning bombardment. If there had been no bombarding, then there would have been some probability of decay, but by a different causal path; one based on the presence of energy state *E*1 and a different law. The two paths are not manifestations of a single path-type. Thus, bombarding was a cause of the decay.

I note that the ED+analysis invokes the counterfactual in [2b], which is essentially counterlegal. It implies the counterlegal below since in supposing ¬O(being-in-*E*1–3-N) and then supposing ¬O(bombard), we are envisioning a situation in which *b* is in no energy state at all, which is physically impossible:

(3) If *p* had been in neither *E*1 nor *E*2, nor *E*3, and so on, it would not have decayed.

I have already commented on counterlegals in this analysis at the end of Section 1. (3) is true because decay is an event that requires a prior energy state in the decayed atom.

We tell a very similar story for the case of spontaneous decay without bombardment. In the simple case we can represent the causal path thus:

[1] O(being-in-*E*1) and O(decay).
[2] There are events *f* – in this case, being-in-*E*2-N – such that:
a. (¬O(being-in-*E*2-N) > (O(being-in-*E*1)) & O(decay)))
b. (¬O(being-in-*E*2-N) > (¬O(being-in-*E*1) > ¬O(decay))).
[3] No non-obtaining *g* is such that:
(¬O(being-in-*E*2-N) > O(*g*)) and:
((¬O(being-in-*E*2-N) & ¬O(*g*)) > ¬O(decay)).

We are now in a position to explain the peculiar form of suppositions that have to be introduced to recover the dependency in this case, discussed at the end of Section 1. If we had simply considered supposition of the form ¬O(being-in-*E*1) then we would not necessarily get a dependency (¬O(being-in-*E*1) > ¬O(decay)), since the condition (being-in-*E*1) might come about by virtue of *b*'s being in the state *E*2, or *E*3, and so on. But by adding the extra antecedent, we restrict the supposition ruling out alternative energy states for *b* and so, through what is essentially a counterlegal, we get a hypothetical dependency. ED+ also certifies this as a real causation, for the actual path is a manifestation of the disposition *D*[decay-by-*x*], where *x* in this case is fixed by the half-life law for the unbombarded atom.[26]

5 Queries

That completes the analysis. There may be some doubts. First, there is the worry that the ED+ account, in its application to indeterminism, is beholden to the physics in an unacceptable way. For example, will there always be energy states with which to construct processes? This query is mistaken since I have argued that processes can be constructed from chance-properties, propensities, themselves, and that without some non-gappy process, we have no grounds to attribute causation just on the basis of chance-raising. Another objection is that ED+ leaves causation without any close connection with experimental situations. Not so! Manipulation in the lab, shooting photons at atoms, shows that bombarding atoms with photons raises the chance of decay. One way of explaining that chance-raising is to hypothesize that a structure as in Figure 7.2 is in place. The information in Figure 7.2 amounts to this: (i) the counterfactual information expressing the condition that, in each case of bombarding and decay, without bombarding a certain path-disposition would not have been manifested, as described in the last section; (ii) the counterfactual information that if there had been no bombarding, there would still have been some chance of decay. Point (i) entitles us to the conclusion that there is causation in the offing. So causal commitment is related to experiment.

A second general issue concerns ED+'s extensional adequacy. I have argued (Barker 2003a) that an outstanding problem for counterfactual analyses of causation is the problem of effects: the problem of determining the right causal order between events linked by causal processes. ED+ does not deal with these problems.[27]

A final issue is that the ED+ account depends heavily on counterfactual embedding, but embedding is not understood very well; there is no good semantics of embedded counterfactuals. My aim here is not to provide such a semantics, but offer a motivation for it. Namely, that by appeal to embedding we can explain both deterministic and indeterministic causation within the framework of one counterfactual analysis.

Notes

1 See Lewis (1986: 175–84), Dowe (2000b: 26–8).

2 This reasoning here is not totally removed from the following: the chance-raising of *p*'s bombardment on *b*'s decay only entitles us to conclude that it was *likely* to some degree that *p*'s bombardment caused *b*'s decay, since, given there is a residual chance of decay without bombardment, there is consequently some likelihood that the actual cause of decay was not bombardment, but a spontaneous event. Lewis (1986: 180) takes himself to have refuted this line of argument. According to Lewis this response assumes that: (i) the bombardment caused *b*'s decay if and only if had the bombardment not occurred, decay would not have occurred; and (ii) one of the following counterfactuals must be true with most likelihood being assigned to (b): (a) *b* would have decayed had it not been bombarded; (b) *b* would not have decayed had it not been bombarded.

3 See Schaffer (2001a) for an account of causal processes in terms of law-subsumption, and Dowe (2001) for one in terms of conserved quantities. A process will not be sufficient for causation by any means, we are only concerned with necessity at this point. The counterfactual analysis of causation to be given below does not depend upon appealing to causal processes, as such.

4 This conclusion would hold even in the face of evidence in the lab in which we manipulate atoms of *b*'s kind by bombarding them with photons. When we bombard them there is a high chance of emission, much higher than without. How can we say there is all this regular chance-raising without causation? The answer is that this is chance-raising in the case of types, not tokens. All we could conclude is that in particular cases, the chances of causation was higher. The issue here is that (particular case) chance-raising is a good indicator of causation but does not in itself constitute causation.

5 Noordhof (1998a: 460) envisions this kind of chance set-up.

6 In the case of Figure 7.1 there is no non-gappy process defined in term of propensities.

7 For example, Helen Beebee (1997) takes this line following suggestions in Dowe (2001).

8 Some probabilistic, counterfactual analyses of causation attempt to avoid reference to processes – for example, Noordhof's (1999) analysis. But I see them as trying to capture the same effect as is gained by explicit reference to processes. (Reference to processes is unavoidable in the semantics for counterfactuals since they are required to determine the character of miracles.)

9 There are some theorists, such as Lewis (1986: 241–69), who might deny that the state of affairs comprising an object's instantiating a propensity could be a cause because such a state of affairs fails to meet the conditions of admissibility on causal relata. On such criteria, such states of affairs could not be parts of causal processes either. The whole line of this paper is against any such criteria on causal relata.

10 I note that there is no issue here of preemption, which might explain why, on the present understanding of chance-raising, *b*'s being in $E1$ fails to be a chance-raiser of decay.

11 There are also good reasons not to introduce chance-raising. How much chance-raising do we need to get causation? Suppose that certain atoms, left unaffected, have a half-life of 100 minutes; so within 100 minutes, 50%, on average, decay. It might have been that

bombarding such atoms with photons produced one of the results: 50.001% … 51% … 70% … 99% decay within an hour. But amongst these possibilities what is the cut-off point for causation? Why wouldn't 50.001% be sufficient? In this case, there is a process linking the event of bombardment to decay. Bombardment brings about a change of energy state in the atoms, which is responsible for decay. If so, why isn't the bombardment a cause since it is responsible for the state that makes the chance outcome possible? If this is so, then chance-raising by any amount is sufficient for causation if a process of the right kind is in place.

12 O(…) is shorthand for … occurs or obtains. > is the counterfactual conditional connective. Note that with respect to condition [2], there may be more than one *f*-event.

13 ED is a descendent of Lewis's quasi-dependency account (see Lewis 1986: 205–7). *e* quasi-depends upon *c* if and only if a causal process links *c* and *e* and in most physically possible worlds in which *c–e* obtains, *e* depends upon *c*. The account has the regrettable feature of vagueness, about dependency in *most worlds*. See Dowe (2001) for a critique of the account. ED finds a cousin account in the work of Ramachandran (1997, 1998) and Ganeri, Noordhof and Ramachandaran (1996, 1998). ED also finds inspiration in the work of Pearl (2000). A simpler kind of counterfactual account is a *holding-fixed* account. *C* causes *e* iff holding fixed some condition *H* obtaining under the circumstances, had *c* not occurred, *e* would not have occurred. This kind of account immediately ushers in backwards causation. Say sodium and water mixed produce an explosion. Then, holding fixed that there was sodium present, if there had been no explosion there would not have been any water present. So, the explosion caused the presence of water. On the ED account there is no causation since Condition [1] is not met.

14 In conceiving of the fusion of the firing of *F*1 to *F*4 not having occurred, we really need to conceive of all of its parts not occurring. But this cannot be a sufficient condition for causation, since if it were, all sorts of causally irrelevant events would be counted as contributory causes.

15 On Paul's (1998b) analysis hasteners are causes. For Paul (1998) *c* causes *e* iff had *c* not occurred *e* would not have occurred or it would have occurred later. The second condition is met in the window case and so the opening of the window caused its own smashing. Lewis (2000) does so likewise.

16 There is no chain-wise dependency, for example.

17 Take the case where there is a third spell-caster, Mandrake, who does not cast, but might if Morgana had not cast. In this case, the counterfactual [3b] is false – *if Merlin had not cast, and Morgana had not been first, then there would have been no frogification* – since the second conjunct of the antecedent could have been made true by Mandrake's casting before Morgana. Condition [3] is then not met. This problem is solved by including the event, *Mandrake does not cast*, as one of the conditions *f* being assumed not to obtain.

18 This assumes that the process is discrete. So as not to assume discreteness: a causal path is one such that there is some finite set of events which are parts of the process, and such that members of each successive pair are linked by causation.

19 Thus, the bomb being present caused Jane to see it. Her seeing it caused her to move away. Her moving away caused her not being present where the bomb exploded. And her not being present where the bomb exploded was a cause of her healthy state moments later, which in turn is a remote cause of her health the next day. By transitivity, the bomb was a cause of her health the next day.

20 See McDermott (1995) for the well-known dog-bite case that instigated a lot of the recent discussion of transitivity.

21 In the bomb case, there is, by ED, a causal-path from the presence of the bomb to Jane's healthy state the next day, since, [1] the bomb was present and Jane was healthy the next day; [2] had none of the alternative paths to health occurred, then the bomb would still

have been present and Jane healthy the next day. Moreover, had these alternatives not occurred, then had the bomb not been present, Jane would not have been healthy; and [3] this causal process was complete.

22 Obviously these are not fundamental laws: they are rather law-based generalities. For convenience, I use the term *law* in the main text.

23 I take the example from a talk by Chris Hitchcock given at the conference on causation and explanation in the social sciences in Ghent, 2002.

24 Unfortunately, space constraints prevent a comparison of the approach to causation and apparent transitivity failure developed here with the approaches in Hall (2000), Hitchcock (2001a) and Yablo (2002). Superficially ED+ bears some resemblance to Schaffer's (2001a) account according to which a cause *c* increases the probability of a process to *e*, of which it may be a part. The difference is that Schaffer's account deals with: (i) the chance-raising of an actual token process; (ii) chance-raising; and (iii) processes understood as continuous physical processes. It differs from ED+ on intransitivity cases; as far as I can see Schaffer's account entails, in the train-case, that Billy's switching caused the train to arrive at the terminus.

25 Here (¬O(being-in-*E*1–3-*N*) is the condition that α is not in the state *E*1, or *E*3 to *E*N, where the latter are possible energy states of an atom of *b*'s type to be in.

26 Note furthermore, that ED+ implies that no matter what degree bombardment raises the chance of decay, bombardment causes decay (see note 11 on p. 135). If bombardment changes the energy state, then that is enough to say it is a cause.

27 I have argued elsewhere that some counterfactuals presuppose causation. I set up a dilemma for counterfactual theories in Barker (2003b) to the effect that counterfactual analyses either must use these counterfactuals, and so be circular as reductive accounts of causation; or, be extensionally inadequate, and fail to explain these cases of causation. ED+ seems to neatly sidestep this problem.

Chapter 8

Routes, processes and chance-lowering causes[1]

Christopher Hitchcock

1 Introduction

Causes often influence their effects via multiple routes. Moderate alcohol consumption can raise the level of HDL ('good') cholesterol, which in turn reduces the risk of heart disease. Unfortunately, moderate alcohol consumption can also increase the level of homocysteine, which in turn increases the risk of heart disease. The net or overall effect of alcohol consumption on heart disease will depend upon both of these routes, and no doubt upon many others as well. This is a familiar fact of life for engineers and policy makers, one that often gives rise to unintended consequences. Suppose, for example, that the American Federal Aviation Administration were to institute new regulations requiring that aeroplanes be equipped with some expensive new safety feature. Would this regulation save lives? Not necessarily. Every dollar (or pound, or euro, or yen …) that an airline spends to upgrade its fleet is a dollar that must be recouped in some way, most likely through higher fares. Higher fares may, in turn, persuade some travellers to drive instead of fly – especially on shorter routes. But it is inherently more dangerous to drive a given route than to fly it, so the net effect of the new regulation may cost lives, rather than saving them (Glassner 1999: 188).

Despite the ubiquity of such multiple connections, existing philosophical theories of causation seem to be very poorly equipped to capture this idea. Probabilistic and counterfactual theories tend to be formulated so as to capture only the net effect of a cause upon its effect. Causal process theories of causation have trouble capturing the distinct routes for the same reasons that they have difficulties with prevention and 'causation by disconnection' (see Schaffer 2000a). By contrast, many of the techniques developed in the causal modelling literature – especially the graphical techniques developed by Spirtes, Glymour and Scheines (1993) and Pearl (2000), among others – are particularly apt for capturing multiple connections. In two recent papers (Hitchcock 2001a, 2001b), I argue that it is possible to adapt certain features of these causal modelling approaches within probabilistic and counterfactual theories of causation, and thus capture the notion of a 'causal route' connecting a cause to its effect. With the help of this notion – and, in particular, with the distinction between causal influence along a causal route and 'net' causal influence – it is possible to resolve a number of problems in the theory of causation. In the present paper, I show how this apparatus may be applied to the problem of chance-lowering causes.

My proposal is very similar in spirit to one offered recently by Phil Dowe (1999, this volume), which employs his notion of a causal process (developed in detail in Dowe 2000b). I will take a good deal of care to distinguish my proposal from Dowe's, and to argue for the superiority of the former over the latter. Given the similarity of the two approaches, this may seem like nit-picking. The common insight, however, is one of sufficient power and importance that it is worth getting the details right.

2 On the prospects for reductive analysis

While I will be responding to what is probably the most common objection to probabilistic theories of causation, I do not endorse a *reductive analysis* of causation in terms of probabilities. I am sceptical about the prospects for such a reduction for a number of reasons.

First, there are two broad traditions that attempt to understand causal relations in terms of probabilities. The first, descending from the work of Reichenbach (1956) and Suppes (1970), attempts to analyse causation in terms of conditional probabilities. The second, stemming mainly from the work of David Lewis (1973b, 1986; but see also Mellor 1995) attempts to analyse causation in terms of counterfactuals whose consequents describe single-case chances. According to both approaches, probabilistic theories of causation do not maintain that causes raise the probabilities of their effects *simpliciter*; rather, causes raise the probabilities of their effects *ceteris paribus*, while holding other factors fixed. But one should not hold *all* other factors fixed, and it is implausible that one can specify which factors *are* to be held fixed in purely *acausal* terms. This problem was raised forcefully by Nancy Cartwright (1979), and has been widely recognized by philosophers who work within the first tradition, such as Eells (1991). Cartwright's objections seem to have had little effect on those who attempt to understand causation in terms of counterfactuals, however. In my opinion, this is a historical accident – counterfactual approaches to causation are no more immune to these objections than their cousins.

Second, there is a class of cases that strike me as more troublesome for probabilistic theories than chance-lowering causes (or chance-raising preventers). These are cases where one event raises (or lowers) the probability of another without being causally relevant to it *in any way*. I discuss these cases in detail in Hitchcock (forthcoming); see also Menzies (1989, 1996), Woodward (1990), and Schaffer (2000b).

Finally, I doubt that our intuitive judgements about what causes what correspond to some unique concept. I suspect, rather, that there are a number of distinct causal relations, and that we attend to one or another of them in different contexts. This is not to say that causation is hopelessly subjective or interest-laden. Whether or not an event stands in some particular causal relation to another is fully objective. But we do not consistently refer to any one such relation as 'causation' in all contexts. See Hitchcock (2003) for a detailed defence of this position.

3 The problem of chance-lowering causes

As noted in the previous section, there are at least two broad approaches to causation that can be dubbed 'probabilistic'. Each of these broad approaches can, in turn, be developed in different ways. The framework that I will develop below can be adapted to fit a number of different theories of causation, but for the sake of definiteness, I will work within the framework of Lewis's counterfactual theory (Lewis 1973b, 1986).

Let *c* and *e* be two distinct events that actually occur. We will say that *e counterfactually depends upon c* just in case:

> the actual chance of *e*'s occurrence, *Ch*(*e*), at the time of *c*'s occurrence, is higher than *Ch*(*e*) would have been, at the same time, had *c* not occurred.

Letting the time of *c*'s occurrence be *t*, and the actual chance of *e* at *t* be *x*, this requires that:

> in all of the closest possible worlds where *c* does not occur at *t*, the chance of *e*'s occurrence, as of time *t*, is less than *x*.

These counterfactuals are to be understood as *non-backtracking*, so that causes do not depend counterfactually upon their effects (see Lewis 1979 for details).

Now we take a first stab at formulating a probabilistic theory of causation:

> *c* is a cause of *e*, just in case *e* depends counterfactually upon *c*.

This formulation runs headfirst into the problem of chance-lowering causes. Here is an example that illustrates this problem:

> *Back-up Assassin*. An assassin-in-training is on his first mission. His victim comes into sight. Given the novice's lack of experience, the victim's awkward location, and so on, the assassin-to-be has only a 30% chance of hitting and killing his target (assuming he shoots at all). The trainee is accompanied by a back-up, a trained assassin who has a 70% chance of hitting and killing the victim. However, the back-up has been given orders to shoot only if the novice fails to shoot at all, and not if he shoots but misses. (It is important that the job be done with a single bullet, so as to avoid the appearance of a conspiracy.) In fact, the assassin-in-training does shoot and kill the victim.

In this example, we regard the assassin-in-training's shot as a cause of the victim's death. But his shot *lowered* the probability of death: if the novice had not shot, then the back-up would have, and the victim's chance of death would have been 70% instead of 30%.

Lewis's own solution to this problem involves the postulate that causation is *transitive*: if a causes b, and b causes c, then a causes c as well. Counterfactual dependence is not transitive in general, so Lewis does not identify causation with counterfactual dependence. Rather, Lewis identifies causation with the *ancestral* of counterfactual dependence. This attempted solution is the same as the one that Salmon (1984) calls the strategy of 'successive reconditionalization', which is criticized by Dowe (this volume). I think that there are independent reasons for denying that causation is transitive (Hitchcock 2001a; see also McDermott 1995). Hence another solution to the problem of chance-lowering causes is wanted.

4 Chance-raising along a causal route

I propose to solve the problem of chance-lowering causes by making use of the representative power of graph-theoretic approaches to causal modelling. The formulation that I will develop is deeply indebted to Pearl (2000), although he would disavow the precise formulation given here. Consider the three most significant events that did occur, or might have occurred, in *Back-up Assassin*: the assassin-in-training shoots; the back-up assassin shoots; the victim dies. Call these events a, b and v, respectively. We may express the pattern of counterfactual dependences among these events as follows:

1a. If a were to occur, b would not occur (or $Ch(b) = 0$)
1b. If a were not to occur, b would occur (or $Ch(b) = 1$)
2a. If neither a nor b were to occur, v would not occur ($Ch(v) = 0$)
2b. If a were to occur but b not to occur, then $Ch(v) = 0.3$
2c. If a were not to occur but b were to occur, then $Ch(v) = 0.7$
2d. If both a and b were to occur, then $Ch(v) = 0.79$

We could tell the story in such a way that the consequents of 1a, 1b and 2a are genuinely chancy, but that adds unnecessary complications.[2]

It is possible to represent this system of counterfactuals very elegantly. Let us introduce *random variables* A, B and V. A takes the value 1 just in case a occurs, and takes the value 0 otherwise. The other two variables are to be interpreted analogously. Then we may express the counterfactuals 1a through 2d as two equations:

(1) $B = 1 - A$

(2) $Ch(V = 1) = 0.3A + 0.7B - 0.21AB$

These equations function like ordinary algebraic equations in some respects, but not in others. In particular, they can be used to evaluate counterfactuals not explicitly listed in 1a through 2d. For example, the assassin-in-training actually shot, that is $A = 1$. Substituting, Equations 1 and 2 tell us that $B = 0$ (the back-up did not shoot), and that the chance of the victim's death was equal to 0.3. Similarly, we can discover what would have happened had $A = 0$; we get $B = 1$ and $Ch(V = 1) = 0.7$.[3] Now let us suppose that the back-up had shot. Since the relevant counterfactual does not backtrack, we do not want to infer that the novice did not shoot. Thus we do not simply substitute $B = 1$ into the left-hand side of Equation 1 and solve for A. Rather, we *replace* Equation 1 with the new equation $B = 1$. We substitute values when a variable appears on the right-hand side of an equation, but we replace the entire equation when the variable appears on the left. Thus Equation 1 does not carry the same counterfactual information as its algebraic equivalent $A = 1 – B$. For further explanation of this convention, see Hitchcock (2001a) and especially Pearl (2000).

The equations that we use to represent a particular situation must be written in a certain minimal normal form. We could, in principle, expand the counterfactual 1a into the following:

1a' If *a* were to occur and *v* were to occur, then *b* would not occur
1a" If *a* were to occur and *v* were not to occur, then *b* would not occur

and analogously for 1b. In much the same way, we could re-write equation 1 as follows:

(1') $B = 1 – A + 0V$

Since the relevant counterfactuals do not backtrack, the inclusion of information about the victim's fate in the antecedent of counterfactual 1 makes no difference whatsoever to the consequent. When equations are written in the appropriate normal form, such irrelevant variables are excluded as arguments. This is because we want the *form* of these equations to convey qualitative information about which variables depend upon which others.

This qualitative information can be neatly represented in a *directed graph*. The variables that occur within the equations representing a situation correspond to the nodes of the graph. In our example, these nodes will be labelled A, B and V. An arrow is drawn from one node to another just in case the former node corresponds to a variable that figures in the equation that determines the value of the second variable. Equation 1 therefore requires that an arrow be drawn from A to B, while Equation 2 entails that arrows are to be drawn from A and B to V. The directed graph is shown in Figure 8.1. (Again, for a more detailed discussion of these procedures, see Hitchcock (2001a) and especially Pearl (2000).)

The directed graph shows in a very clear and heuristically powerful way that the novice's shot affects the victim's prospects for death via two distinct *causal routes*. These two routes correspond to the two directed paths that connect A to V in Figure 8.1. Intuitively, the 'direct' route from A to V represents the influence that the trainee exerts by his choice of whether or not to send a bullet on its way towards the victim. The 'indirect' route, which runs through B, represents the effect that the

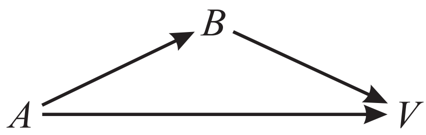

Figure 8.1

trainee's choice has in virtue of its effect on whether the back-up assassin shoots. Note the counterfactual facts that determine that an arrow should be drawn directly from A to V: whether or not the victim dies depends counterfactually upon whether or not the novice assassin shoots, *even when we hold fixed whether the back-up assassin shoots.* Because the consequent of counterfactual 2b is different from that of 2a, and the consequent of 2d is different from that of 2c, the novice's shot is relevant to the victim's death independently of its effect on what the back-up does.[4] In comparing these two pairs of counterfactuals, we 'freeze out' the indirect effect of A on V through B by stipulating a value of B in the antecedents of the counterfactuals. Elsewhere (2001a), I refer to such counterfactuals as 'explicitly non-foretracking' or ENF counterfactuals, since the effect of counterfactually varying A is not allowed to 'foretrack' to V.

The routes depicted in Figure 8.1 correspond to paths of *possible* influence from A to V: whether or not A *actually* affects V along either route will depend upon the *actual* values of the variables A and B. Similarly, the arrows say nothing about the *nature* of A's effect on V – they tell us nothing, for example, about whether a tends to cause or to prevent v along a given route. To know whether a counts as an actual cause of v or not, we need to look at a restricted set of the counterfactuals captured by Equations 1 and 2. In fact, a occurred ($A = 1$), b did not occur ($B = 0$), and v did occur ($V = 1$). Thus, of the counterfactuals 2a through 2d, 2b is the one whose antecedent corresponds to the actual state of affairs. Strictly speaking, 2b is not a *counterfactual*, but rather a subjunctive conditional with a true antecedent. To assess the actual impact of the novice's shot, along the direct causal route, we must determine what the chance of v would have been had the novice not shot, while still stipulating that the back-up assassin did not shoot. The answer is given in the consequent of 2a. Holding fixed that the back-up assassin did not shoot, the chance of v would have been lower (0 instead of 0.3) had the assassin-in-training not shot. It is in this sense that the novice's shot can be said to increase the chance of the victim's death. It does not increase the chance of the victim's death *overall*, but its effect along the direct route is to increase the chance of death.

We are now in a position to state the solution to the problem of chance-lowering causes. Let c and e be distinct occurrent events, which correspond to values of the variables C and E respectively. Then c is a cause of e just in case c raises the probability of e along some causal route r from C to E; that is, just in case the actual chance of e's occurrence[5] is greater than the chance of e's occurrence in the closest possible world(s) where c does not occur, but where all the actual events[6] that lie between c and e along *other* causal routes from c to e nonetheless do occur.[7]

5 Dowe's path-specific solution

Dowe (1999; this volume) presents an account of chance-lowering causes that bears a strong resemblance to that sketched above. Dowe's account makes use of his notion of causal process. According to Dowe, a causal process is a world-line of some object that possesses some value of a conserved quantity (see Dowe (2000b) for details). Dowe delineates the distinct paths that connect a cause to its effect in terms of causal processes; he is aware, however, that a simple-minded identification of causal paths with causal processes will not work. For example, Dowe would agree that there are two paths[8] connecting the trainee's shot with the victim's death in *Back-up Assassin*. One of the paths is constituted by a causal process: the novice's bullet. The second, however, is not constituted by a causal process *per se*, but rather by a *potential* causal process.[9] After the assassin-in-training fires his weapon, there are causal processes that carry information about this event to the back-up assassin – photons or sound waves, perhaps. But these processes do not continue on to the victim,[10] nor do they initiate some new causal process that connects the back-up assassin to the victim. Rather, the processes originating from the novice's gunshot *interrupts* a process that *would have* connected the back-up assassin with the victim. This second path thus has the character of a *prevention*, a species of what Dowe (2000b, ch. 6) calls *causation**.

The novice assassin's shot raises the chance of the victim's death *via* the first path 'in itself' (Dowe, this volume: 35). This relation between a and v is determined by actual and counterfactual chances of v, not in the actual world, but in 'the closest worlds in which that path is the only path between' a and v (ibid.: 35). That is, in 'worlds in which there was no way the victim could have died except by [the novice's] bullet' (ibid.: 35), the chance of the victim's death would be higher were the trainee to shoot than it would be if he were to refrain from shooting. It should now be clear that Dowe's proposal is really strikingly similar to my own, for one class of possible 'worlds in which there was no way the victim could have died except by [the novice's] bullet' will be the class of worlds in which the back-up assassin refrains from firing regardless of the trainee's action. That is, the counterfactuals 2a and 2b that are to be compared on my account are precisely the sorts of counterfactuals that are to be compared on Dowe's account as well.

The two accounts differ with respect to how the distinct paths that connect cause and effect are to be delineated. According to my account, this is determined entirely by the structure of the counterfactuals that characterize a given scenario. According to Dowe's account, the decomposition of influence into distinct paths is determined by the actual and potential *causal processes* that link the events in question. To make this distinction vivid, note that on my analysis Figure 8.1 would accurately characterize the structure of *Back-up Assassin*, even if the assassins' guns killed by some sort of unmediated action-at-a-distance. All that is required is that the counterfactuals 1a–1b, and 2a–2d be true.

The complaint that I will be levelling against Dowe's account is that it does not go far enough in telling us when two paths are genuinely distinct. Determining when causal paths are genuinely distinct is essential if genuine cases of probability-lowering causation are to be distinguished from imposters.

6 Delineating causal routes

In order to present this complaint more precisely, it will be helpful to return to the positive account developed in Section 4 above. There we introduced two variables: A, which takes the value 1 or 0 depending upon whether or not the novice shoots; and B, which takes the value 1 or 0 depending upon whether or not the back-up assassin shoots. But was this really necessary? As the scenario was described, exactly one of the two assassins would shoot.[11] So why not introduce a single variable, S, which takes the value 1 if the trainee shoots (that is, if a occurs), and takes the value 0 if the back-up shoots (b occurs)? Then we would arrive at a much simpler description of the scenario, characterized by the following counterfactuals:

3a If a were to occur, then $Ch(v) = 0.3$
3b If b were to occur, then $Ch(v) = 0.7$

with corresponding equation:

(3) $Ch(V = 1) = 0.3 + 0.4(1 - S)$

If we represent this graphically, we will have two nodes, S and V, with an arrow running from S to V (see Figure 8.2). In this representation, there is only one route from S to V. Along this route, a *lowers* the chance of v, thus reinstating the problem of chance-lowering causes. If my solution is to succeed, some clarification is in order: there must be a principled reason for taking Equations 1 and 2 to be the correct way to characterize the scenario, and for taking Equation 3 to be inappropriate.

In order to address this question, we must ask what it means to represent two events (such as a and b) as different values of the same variable, or as values of different variables. When we represent two events as different values of the same variable, we are representing those events as *mutually exclusive*. A variable is a function (over possible worlds, if you like), and hence it must be single-valued. Moreover, the two events will be exclusive, regardless of the equations that represent the system. In particular, the exclusion of one event by the other will not correspond to any of the arrows that figure in the corresponding graph. What this suggests is that the relevant form of exclusion is not *causal*, but logical, conceptual or metaphysical. In our example, the novice's shot *prevents* the back-up from shooting. This is a causal relationship between the two events, corresponding to the arrow from A to B in Figure 8.1. This causal relationship is concealed in Figure 8.2. We may thus offer

$$S \longrightarrow V$$

Figure 8.2

the following rule of thumb: two events are to be represented as values of different variables if: (a) they are not mutually exclusive; or (b) they are mutually exclusive, but the exclusion is causal – one event *prevents* the other from occurring.

This rule is far from satisfactory: it appeals to causal facts, and the current project is to recover causal facts from algebraic and graphical structures that are defined in terms of counterfactuals only.[12] Nonetheless, the foregoing considerations point us in the right direction. We noted in Section 3 that within counterfactual theories of causation, causation can only hold between *distinct* events. The reason for the restriction to distinct events is familiar by now: we don't want to say that my raising my arm caused my arm to go up, that my saying 'hello' caused me to say 'hello' loudly, that my stroll caused my first fifty steps, and so on. In each case, there is counterfactual dependence – if I hadn't raised my arm, it wouldn't have gone up – but not causation. This problem was posed by Kim (1973). The solution is to note that in each case the events in question fail to be distinct; Lewis (1986) develops a detailed account of event distinctness. What ought to be apparent (but is never discussed) is that the same problem can arise for cases of *prevention*: we also want to avoid saying that my raising my arm prevented my arm from going down, that my saying 'hello' prevented me from remaining silent, and so on. Although we have counterfactual dependence in each of these cases, we fail to have genuine prevention. The reason again involves a failure of distinctness: my raising my arm is not distinct from the failure of my arm to go down. Some may find this particular formulation jarring: how can an occurrent event fail to be distinct from an event-absence? No matter, we can just introduce some different terminology: my raising my arm and my arm's going down are *contrary* events, where contrariety is to be explicated in terms of logical and spatiotemporal exclusion in the spirit of Lewis (1986). The key point here is that a counterfactual theory of causation is *already* committed to a notion of contrary events. We may now use this notion to re-formulate our rule: two events are to be represented as different values of the same variable if they are contrary; as values of different variables if they are distinct but not contrary.[13] In *Back-up Assassin*, the novice's shot and the back-up's shot are not contraries; hence they must be represented as values of different variables, and the one could (in principle) cause or prevent the other.

Let us illustrate the rule with a further example, taken from Cartwright (1979).

> *Weed.* A weed in a garden is sprayed with a defoliant. This decreases the chance that the weed will survive from 0.7 to 0.3. Nonetheless, the weed survives.

Intuitively, spraying the weed with defoliant did not cause it to survive. Note, however, that the probabilities are identical to those in *Back-up Assassin*. What is the difference between the two cases, such that we regard one to be a case of causation, the other not? In order to answer this question we must look at the equation(s) and graph that characterize *Weed*.

Here is a natural attempt to provide a representation: Let S' be a variable that takes the value 1 if the weed is sprayed, 0 if it is left alone; and let V' take the value 1 if the weed survives, 0 if it dies. Then we can represent the situation using Equation 3 and Figure 8.2, adding primes to the variable names.

Is this the correct representation? Or should the representation look more like Figure 8.1 and equations 1–2? In order for the latter to be correct, we would have to include one more variable in the model. We might try to do this in the following way: let A' take the value 1 or 0 according to whether the weed is sprayed or not, B' take the value 1 or 0 according to whether the weed is left alone or not. Using these variables, the analogues of equations 1 and 2 seem to capture the relevant facts. But this representation clearly violates our rule: $A' = 0$ and $B' = 1$ represent events that are *not* distinct, while $A' = 1$ and $B' = 1$ represent events that *are* contrary.

Alternately, we might note that spraying the plant affects its chances of survival by affecting its state of health shortly after the spraying. Let S' and V' be defined as above, and let $H = 1$ or 0 according to whether the plant is healthy or not, one day after being sprayed. In order to make the case parallel to that of *Back-up Assassin*, suppose that the weed will be healthy just in case it is not sprayed. That is, suppose that the equation expressing the relationship between S' and H is:

(4) $H = 1 - S'$

Now, however, there are (at least) two different ways in which we can write the second equation so as to preserve the appropriate probabilities. We could write it on the model of equation 2:

(5) $Ch(V' = 1) = 0.3S' + 0.7H - 0.21S'H$

or we could write it more simply:

(5') $Ch(V' = 1) = 0.3 + 0.4H$

Both entail that the plant has a 0.3 chance of surviving if it is sprayed, 0.7 if it is not. There is an important difference, however: Equation 5 implies that the plant's chance of survival depends upon whether or not it is sprayed, even when its later state of health is held fixed; Equation 5' does not. This is an empirical matter, not settled uniquely by the description of the scenario. Nonetheless, Equation 5' seems vastly more plausible: spraying the weed affects its chance of survival only by affecting its subsequent health. If the plant were sprayed, but were (miraculously) healthy the next day, its chance of survival would be 0.7, not 0.79. The fact that it was sprayed, in addition to being healthy, would not give it an *extra* chance to survive – one that it would not have had if had not been sprayed. To slightly abuse some familiar terminology: the state of the plant's health *screens off* spraying from survival.[14]

The graphical representation of *Weed*, as characterized by Equations 4 and 5', is shown in Figure 8.3. This figure shows clearly that there is only one route from the

$$S' \longrightarrow H \longrightarrow V'$$

Figure 8.3

spraying to the plant's survival. It is possible to *interpolate* variables along this route, but doing so does not create distinct routes from S' to V'. The causal structure depicted in Figure 8.2 can be *embellished*, but not fundamentally altered. Since there is only one route from spraying to survival, and spraying lowers the probability of survival, it follows that spraying must lower the probability of survival along this route. There is no event, lying off this route, such that when we hold fixed the occurrence of this event, spraying increases the plant's chance of survival.[15]

In some cases where one event lowers the chance of another, we consider the first to be a cause of the second. *Back-up Assassin* describes one such case. In other cases where the probabilities are the same, such as in *Weed*, we do not consider the first event to be a cause of the second. A satisfactory account of chance-lowering causes must be able to discriminate between these two sorts of cases, and I have shown how my account of chance-raising along a causal route can do this.

7 Delineating causal paths: Dowe's account

I turn now to the question of whether Dowe's path-specific solution to the problem of chance-lowering causes can make the relevant discriminations. The worry is that this account will prove too much: that it will yield chance-lowering causation not only in cases such as *Back-up Assassin*, but also in cases like *Weed*. More specifically, the worry is that Dowe's account will yield the result that in *Weed*, just as in *Back-up Assassin*, we have a case where:

> 1. A cause and its effect [are] linked along more that one path …
> 2. Two paths between a cause and its effect are 'opposed' … [O]ne of the paths is a causing path, and the other a preventing path …
> 4. … [V]ia the successful path 'in itself' the cause *c* raises the chance of *e*.
>
> (Dowe, this volume: 35)

In *Weed*, there is an actual causal process connecting the spraying of the weed with its later survival. This process includes intermediate stages in which the plant is in a sickly state. Is there another path as well? Spraying the plant *prevents* the occurrence of a connecting process consisting of stages in which the plant is healthy. But is the actual causal process distinct from this preventing path? It is at this point where Dowe's account fails to supply the necessary details.

Although Dowe is not entirely clear on this point, the idea seems to be that the actual causal process is distinct from the preventing path just in case the actual causal process is in some appropriate way *different* from the alternative causal process that would have occurred. Thus in *Weed*, the question is whether the actual causal process – consisting of sickly plant stages – is different from the causal process that would have resulted if the plant had not been sprayed. Dowe *is* entirely clear that more than a difference in spatiotemporal location is required. It is easy enough to invent variants on *Back-up Assassin* where the back-up process follows that same spatiotemporal trajectory as the novice assassin's bullet; and variants on *Weed* where the plant changes location depending upon whether it is sprayed. One natural suggestion would be that the two processes are different just in case they involve the transmission of different values of the relevant conserved quantities. But this proposal is unlikely to give us the answer we want: it seems that spraying the plant *would* affect the amount of charge, energy, and so on transmitted by the plant, and that it would rearrange the distribution of these quantities among parts of the plant. It thus appears that Dowe is committed to saying that there are indeed two opposing paths connecting the weed's spraying to its later survival.

Some of Dowe's more informal comments further suggest that his solution is committed to treating *Weed* as a case of chance-lowering causation. He writes:

> The path-specific chance relation between *c* and *e* is given by the chance relation between *c* and *e* in the closest worlds in which that path is the only path between *c* and *e*. Take worlds in which there was no way the victim could have died except by [the novice's] bullet … in [those] world[s] $ch_C(e) > ch_{\sim C}(e)$.
>
> (Dowe, this volume: 35)

By parity of reasoning, when examining *Weed*, we should look at 'worlds in which there was no way the [weed] could have [survived] except by [the sickly process]'. We might imagine worlds in which some agent is prepared to kill the plant just in case it is healthy; or worlds in which the laws of plant physiology are different, so that states that are healthy in our world are lethal in those worlds. In such worlds, spraying the weed does indeed increase its chance of survival: it prevents the plant from continuing in a 'healthy' state that would guarantee its death. Thus it seems that the spraying does indeed raise the chance of survival via the sickly process 'in itself'.

Note that the issue is not whether my account or Dowe's better captures the intuitive notion of a causal path or route. There may be a perfectly coherent concept of causal path wherein there are two paths from spraying to survival – one via a healthy intermediate state, the other via a sickly one. But in order for the concept of a causal path to do the work it is supposed to do in Dowe's account of chance-lowering causes, there has to be only one path from the spraying to the survival. My account can yield this verdict; Dowe's account cannot.

I have subjected Dowe's account of chance-lowering causes to a rather harsh critique. This is not at all because I think that the account is misguided. To the contrary, I wholeheartedly agree with the general outlines of Dowe's solution. It is precisely because the central idea – that causes may be connected to their effects via multiple paths – is so powerful, that it is appropriate to press hard for details.

Notes

1 For comments and discussion, thanks go to Phil Dowe, Jonathan Schaffer and James Woodward.
2 For those readers who are interested, I will include a brief discussion of some of these complications in the notes.
3 Both of these computations are more complicated if the connection between A and B is chancy. Consider the hypothetical case where $A = 0$. Let us write a' to represent the absence of event a. Now we must distinguish between the chance conferred upon v by a', $\mathrm{Ch}_{a'}(v)$, and the chance conferred upon v by a' together with b, or $Ch_{a'b}(v)$. (Mellor 1995 is very clear on the need to specify the facts that are conferring a chance upon an event.) We then calculate $Ch_{a'}(v) = Ch_{a'b}(v)Ch_{a'}(b) + Ch_{a'b'}(v)Ch_{a'}(b')$. This formula is more familiar in the notation of conditional probability: $P(v/a') = P(v/a'b)P(b|a') + P(v/a'b')P(b'|a')$. For further discussion, see Spirtes, Glymour and Scheines (1993) and Pearl (2000, especially chapter 2).
4 Note that an arrow would be drawn from A to V if only one of these pairs of consequents was different: V must depend counterfactually upon A while holding B fixed at *some* value.
5 More precisely, the actual chance of e's occurrence in virtue of the occurrence of *c and* of various specified events that lie between c and *e along other causal routes from c to e.* This must be stipulated explicitly, since the occurrence of these other events may not be determined at the time of c's occurrence.
6 The phrase 'all the actual events' requires some elaboration. First, the word 'event' here really means 'value of a causal variable'. Thus the back-up's failure to shoot (corresponding to $B = 0$) counts as an 'event' in the relevant sense, even though it might be more proper to describe this as the absence of an event, rather than an event in its own right. Second, the word 'all' is intended to take care of a potential problem not explicitly discussed in the text. In the example discussed, the 'indirect' route from A to V is deterministic. In order to 'freeze out' a deterministic route, it suffices to hold fixed the occurrence of any one event along the route in question. If the route has a finite number of chance points – if for instance, it is a chancy matter whether the back-up will pull the trigger, and also whether her gun will function properly, but all other processes are deterministic – then it suffices to hold fixed the occurrence of the last chancy event. If there are an infinite number of chancy events in the chain (and hence no last chancy event), then it becomes necessary to hold fixed *all* the events along the route in question, except for some initial segment. In each case, it suffices to hold fixed all the events that lie along a route, but in no case is this strictly necessary.
7 Those readers who are familiar with the proposal advanced in Hitchcock (1995a) – bless their souls! – will recognize that the proposal of the current paper is at odds with that advanced earlier. I now think that the central example of that paper, an example involving Sherlock Holmes adapted from Good (1961), should be analysed along the lines suggested here. There are, however, other examples for which my original proposal still applies. I argue (forthcoming) that Deborah Rosen's well-known golf ball example (Rosen 1978) should be analysed along the lines of Hitchcock (1995a). I thus think that there are two separate solutions to the problem of chance-lowering causes, where the present paper confines itself to the solution that has not been discussed at length in earlier publications. Which solution applies in a given context will depend upon whether the different potential causes of the effect in question are to be thought of as different values of *one* variable (in which case the earlier solution applies) or as values of different variables (as in the example of the novice assassin). This is an issue I take up below; see also Hitchcock (2001a).
8 Dowe uses 'path' where I use 'route'. I take it that we are trying to capture the same concept, but will adopt the terminological convention of using 'path' when discussing

Dowe's view, and 'route' when discussing my own.

9 Dowe's usage is not entirely consistent on this point. Sometimes he speaks of potential paths (for example Dowe 1999: 500) rather than paths consisting of potential processes.

10 Or if they do, they are to be ruled as irrelevant to the victim's death. What distinguishes relevant from irrelevant processes is a thorny issue, one I raise in earlier work (Hitchcock 1995b, 1996). Dowe attempts to address this issue (1999: first part; 2000b: chapter 7).

11 This would not be true if the connection between the trainee's shot and the backup's shot were *chancy*. In that case, few would be tempted to commit the error described in the text.

12 As I mention in Section 2, I am not sanguine about the prospects for a reductive analysis of causation, so I do not find this problem to be devastating. Nonetheless, this sort of circularity is to be avoided if at all possible, and in this case it is possible.

13 I assume that if two events are contrary, then they are distinct. The rule gives no advice regarding events that are not distinct. This is a complex problem, but I would suggest something along the following lines: the values of variables in a model correspond to events that are, relative to a certain grain or level of analysis, atomic. Other events are then represented by conjunctions and disjunctions of values of variables; events will fail to be distinct when they contain overlapping components.

14 The terminology is from Reichenbach (1956). The notion of screening off is normally formulated in terms of conditional probabilities, rather than counterfactual chances.

15 I am not denying that we can tell a story in which there is such an event. Perhaps the smell of the defoliant scares off a rabbit who would otherwise have eaten the plant. But in this sort of case, it would be correct to say that spraying the plant caused it to survive.

Chapter 9

Indeterministic causation and varieties of chance-raising

Murali Ramachandran

The world is *indeterministic* if some actual event might have failed to occur without violation of any actual laws; likewise, it is indeterministic if some event that did not in fact occur might have occurred without violation of any laws. One might also make the point in terms of *chance*: in a deterministic world the chance of an event's occurring, given the prior history of the world, is either 0 or 1, whereas in an indeterministic world the chance may be strictly between 0 and 1.

If we want to allow that there is *causation* even in indeterministic worlds, there is little alternative but to take causation as involving *chance-raising*.[1] In the most basic case, one event, C, is a cause of another, E, because the chance of E's occurring is higher as a result of C's occurrence. There are, however, different ways of cashing out the idea of chance-raising. David Lewis (1986) does so by appeal to *counterfactual* chances, for example, whereas Igal Kvart (1986, see also the chapter in this volume) appeals to actual *conditional chances*.

My concern in this chapter is with the counterfactual approach. Even setting aside the notoriously problematic issue of *preemptive* causation,[2] Lewis's theory actually falls down in simple cases that do not involve preemption or over-determination. I shall focus on these simple cases in developing alternative notions of chance-raising and an accompanying account of causation. Serious problems remain, but I think the approach shows more promise than other counterfactual theories of indeterministic causation on the market.

The test cases we shall be focusing on are best introduced by way of considering Lewis's theory.

1 Lewis's probabilistic account

Case 1: Causal chance-raising chains

Consider the following scenario:

The diagram in Figure 9.1 represents a series of neurons linked by stimulatory axons (depicted by forward arrows); shaded circles represent neurons that have fired; unshaded circles (in later diagrams) represent ones that have not. Suppose the *c*–*e*-process is a very reliable one, but that the neurons have a minute background

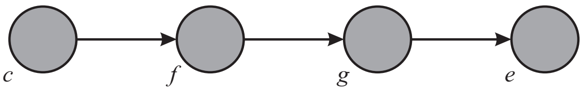

Figure 9.1 Causal chain

chance of firing *un-stimulated*; thus, *e* might have fired even if *g* had not; *g* might have fired if *f* had not, and so on. Intuitively, the firing of *c* (which I shall signify by suffixing an asterisk, thus: *c**) nevertheless is a cause of *e*'s firing (*e**). The reason, presumably, is that the *chance* of *e*'s firing is higher *as a result* of *c*'s firing.

Lewis's (1986) theory develops the idea as follows.

Lewis's Probabilistic Analysis (LPA)

(LP1) For any actual events *a* and *b*, *a raises the chance of b iff* the chance of *b*'s occurring would have been much smaller than it actually is if *a* had not occurred.

(LP2) For any actual events *c* and *e*, *c causes e iff* there is a chain of actual events [c, d_1, … , d_n e] such that each event in the chain raises the chance of the next. (Let's call such a chain of events a *chance-raising-* (or *CR-*)*chain*.)

There are two important points to consider. First, Lewis takes chances to vary over time. So, when we are considering whether an event *a* raises the chance of an event *b*, the actual chance of *b*'s occurring is its chance at the time immediately after *a*'s actual occurrence; and the counterfactual is to concern chance at that same time (Lewis 1986: 176–7). The second point is that for any time *t after* an event *e* has occurred, the chance *at time t* of *e*'s occurring is 1 (this view is made explicit in Lewis 1980: 91).[3]

Returning to Figure 9.1: *c** now comes out as a cause of *e**, as desired, because the chance of *e*'s firing, assessed at time t_c^+ (immediately after *c* fires), will be higher than it would have been if *c* had not fired – that is *c** raises the chance of *e**. However, there are simple counterexamples to the account.

Case 2: Incomplete causal chains

The following example shows that chance-raising is in fact not sufficient for causation.

Suppose here that the stimulatory axons between the neurons are very reliable; *c* and *f* fire, but *g* does not; *e*, however, had a small background chance of firing

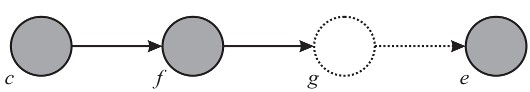

Figure 9.2 Incomplete causal chain

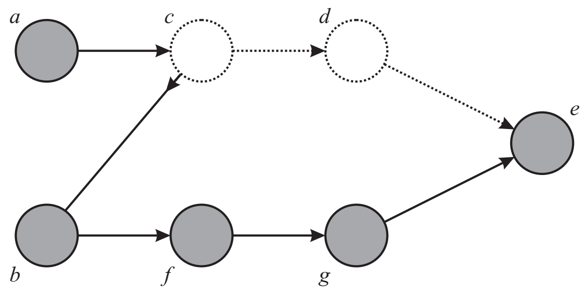

Figure 9.3 Early preemption

regardless, and it does so. In this situation, c^* still raises the chance of e^* – remember, the chance of e^* is evaluated immediately after c^* – but, clearly, c^* does not cause e^*.

Case 3: Transitivity

By Lewis's analysis causation is *transitive*: if there is a CR-chain linking an event *c* to an event *d* and one linking *d* to an event *e*, there *ipso facto* is a CR-chain linking *c* to *e*. But there appear to be straightforward counterexamples to transitivity.

Here's a variation on an example from McDermott (1995: 531 ff.). A dog attacks Singh (event *c*). Singh was due to detonate a bomb the following day, but her nerves have been shattered by the dog-attack. Patel, the only other person qualified to do so, detonates the bomb instead (event *d*). The bomb duly explodes (event *e*). Intuitively, *c* causes *d* and *d* causes *e*, but *c* does *not* cause *e*.

Notice, however, that the dog-attack is not itself a chance-raiser of the explosion. So, it might be thought at this juncture that one could simply dispense with CR-chains, and make it a necessary and sufficient condition of *c*'s causing *e* that *c* raises the chance of *e*, period. This would handle case 1 and the above failure-of-transitivity example. But, in cases of 'early' preemption the cause is *not* a chance-raiser of the effect but *is* linked to it by a CR-chain. For example, consider the scenario in Figure 9.3.

Suppose the stimulatory axons connecting neurons *a*, *c*, *d* and *e* are very reliable whereas the axons connecting *b*, *f*, *g* and *e* are very unreliable, and the *inhibitory* axon (signified by the backwards arrow) from *b* to *c* is very reliable. Event b^* (the firing of *b*) successfully inhibits c^* and the *b*–*e* process runs to completion. Presumably, b^* is a cause of e^*. Yet, b^* in fact *lowers* the chance of *e*'s firing. It would seem we need to appeal to the CR-chain linking b^* and e^* to get the desired verdict that b^* causes e^*.

A dilemma emerges: either we take the existence of a CR-chain to be sufficient for causation, in which case we get the dog-bite coming out as a cause of the explosion, or we don't consider it sufficient, in which case we don't get b^* coming out as a cause of e^* in the Figure 9.3 example.

Case 4: Timely chance-raising

Suppose *c* is a chance-raiser of *e*; *c* and *e* occur, but *e* does not occur within the period of time in which *c*'s occurrence makes a difference; in other words, *e* occurs spontaneously earlier (or later) than any time *c could have* caused it. Lewis's theory still delivers the incorrect verdict that *c* is a cause of *e* (because it takes chance-raising to be sufficient for causation).

An obvious remedy would be to insist a cause of *e* must raise the chance of *e*'s occurring *when it did*. However, we must not take this – *timely-chance-raising* as we might call it – to be sufficient for causation. Suppose, for example, X drinks some water after taking lethal poison and that the water delays her death by a few seconds. X's drinking of the water (event *c*) raises the chance of the death (event *e*) occurring *when it did*; yet, *c* is surely not a *cause* of *e* itself. So, if we are to avoid mere hasteners and delayers of an event *e* coming out as causes of *e*, we should, I suggest, require *both* chance-raising and timely-chance-raising of *e*.[4]

Case 5: Potentially early occurrence of effect

Consider the Figure 9.1 scenario again. Intuitively, *c*'s firing (*c**) causes *f*'s firing (*f**). However, once we allow that *f* might have fired even if *c* had not, it is not unreasonable to also allow that there is a *period of time* over which *f** might have occurred spontaneously. Let us suppose then that *f** might have occurred at the time *c* actually fired, time t_c. So, if *c** had not occurred, *f** might have occurred at time t_c; but if *f** does occur at that time, the chance of *f**'s occurring, *assessed at time* t_c^+ (immediately *after c*'s actual firing), will be 1. Hence, it is not true that the chance of *f**'s occurring would have been much smaller if *c** had not occurred: *c** is not a chance-raiser of *f** in the hypothesized situation. Thus, *c** does not come out as a cause of *f**, contrary to intuition.

One might attempt to defend LPA by maintaining that because the chance of *f*'s firing *spontaneously* is so small, the nearest worlds in which *c* does not fire will be ones in which *f* does *not* fire before t_c^+. But, Lewis himself reckons 'It is fair to discover the appropriate standards of similarity from the counterfactuals they make true, rather than *vice versa*' (1986: 211). And, the counterfactual:

(1) If *c* had not fired, then *f would not* have fired before t_c^+

seems just plain false. Indeed, it sounds no more plausible than the counterfactual:

(2) If *c* had not fired, then *f would not* have fired,

which we have taken to be false *by hypothesis*. So, I do not think this line of defence is credible. Finally, it strikes me that an account that does not have to *rely* on the falsity of (1) in order to solve this problem is to be preferred. I intend to provide such an account.

The foregoing examples motivate the account to be developed here. A limitation of that account – a limitation inherited from LPA – is that it assumes (or, rather, requires) that causes invariably precede their effects. Lewis considers it a virtue of his original (1973) account of deterministic causation that it allows 'backwards' causation, that is, causation of an event by a later event. But his probabilistic account LPA does not allow an event *c* to be a chance-raiser of an earlier event *e* that *might* have occurred even if *c* had not. For the chance of *e*'s occurring, assessed at time t_c^+ will be 1 in the actual world, and *might still have* been 1 *at that time* if *c* had not occurred. The ruling out of backwards chance-raising in such cases thereby precludes backwards causation too.

Now, I do not find the notion of backwards causation incoherent, and I have attempted to accommodate it elsewhere (see Ramachandran forthcoming). But that strategy involved a radically different conception of chance I merely gestured at and which many theorists will find infeasible. In this chapter I shall settle for an account whose key notions of chance and chance-raising are derivable within Lewis's own framework. My aim is not to provide a counterexample-free analysis of causation so much as to see how far one can go with Lewis's original project.

2 Varieties of chance-raising

2.1 Late chance-raising

Let's begin by considering a simple modification to Lewis's account. For any candidate cause and effect, *c* and *e*, Lewis's account assesses the chance of *e*'s occurring at time t_c^+, immediately after *c*'s actual occurrence. Instead, what if we assessed the chance of *e* at the later time t_e^-, the time just before *e*'s actual occurrence, leaving the other definitions of Lewis's theory intact? (We may call the resulting variety of chance-raising *late chance-raising* and Lewis's variety *early chance-raising*.)

This shift in focus to late chance-raising secures the correct verdict in the second test case (the Figure 9.2 example). For, it is settled by time t_e^- that *g* does not fire, that the *c*–*e*-chain is 'broken', as it were; thus, the chance, *at that time* of *e*'s firing is *not* higher as a result of *c*'s having fired.

The proposed manoeuvre also handles the first test case (see Figure 9.1), but the solution here differs in an important respect from Lewis's. LPA delivers the correct verdict that the firing of neuron *c* (*c**) is a cause of the firing of *e* (*e**) because *c** is an early chance-raiser of *e**. However, *c** is *not* a *late* chance-raiser of *e**. By hypothesis, the neurons in this diagram might have fired even if they were not stimulated. It is then feasible that *g* might have fired even if *c* had not; but, so long as *g* fires, the chance at time t_e^- of *e*'s firing will be the same as it is in the actual world; hence, *c** does not come out as a late chance-raiser of *e**. The present account gets the correct conclusion, that *c** causes *e**, not because *c** raises the chance of *e**, but because of the existence of a late-chance-raising chain between *c** and *e**.

If the existence of a CR-chain is taken as sufficient for causation, we are still saddled with the counterexample to transitivity mentioned on p. 154 (the third test case, involving the dog-bite and detonated bomb). The later problem cases, involving the *time* at which the effect occurs or might occur, are also left untouched. So, the present proposal, while an improvement over Lewis's original account – at least as regards the problems we are considering – will not do as it stands.

2.2 Resolving test cases 4 and 5

In test case 4, *c* raises the chance of *e* but *e* in fact occurs (spontaneously) at a time when *c* has no relevant influence. I suggested an obvious remedy for this problem: make it a necessary condition of *c*'s causing *e* that *c* raises *both* the chance of *e*'s occurring *and* the chance of its occurring *when it did*. Test case 5 also has an obvious resolution. The problem here was that *c** fails to come out as a chance-raiser of *f** because *f* might have* occurred spontaneously *before* the time at which its chance of occurring is assessed, at time t_c^+. Clearly, late chance-raising is ruled out as well, because in the same scenario *f** might have occurred before time t_f^-. The solution, I suggest, is to shift our attention from the nearest worlds in which *c** does not occur to the nearest worlds in which *c** does not occur *and in which f* does not occur by time* t_f^-. In such a world, the fact that *f** does not occur by t_f^- is not taken into account when *f**'s chance of occurring is assessed at time t_c^+ or t_f^-: just the laws and prior history of the world of evaluation are relevant. What *is* ensured by our focusing on such worlds is that the chance of *f**'s occurring, assessed at t_c^+ or t_f^-, will never be 1.

Combining the above proposals yields the following analysis of causation.

Defn1: For any actual event *e*, and times t_1 and t_2, let $ch(\text{at } t_1, e)$ be the chance, assessed at t_1, of *e*'s occurring, and $ch(\text{at } t_1, <e, t_2>)$ be the chance, assessed at t_1, of *e*'s occurring at time t_2.

Defn2: For any actual events *c* and *e*, *c* is a *late chance-raiser* of *e iff* $ch(\text{at } t_e^-, e)$ would have been much smaller than it actually is if *c* had not occurred and *e* had not occurred by t_e^-.

Defn3: For any actual events *c* and *e*, *c* is a *timely chance-raiser* of *e iff* $ch(\text{at } t_e^-, <e, t_e>)$ would have been much smaller than it actually is if *c* had not occurred.[5]

Defn4: For any actual events *c* and *e*, *c directly causes e iff c* is a late chance-raiser of *e* and a timely chance-raiser of *e*.

> *Analysis 1*
> For any actual events *c* and *e*, *c causes e iff* there is a chain of actual events $[c, d_1, \ldots, d_n, e]$ such that each event in the chain *directly causes* the next. (I shall call such a chain a *direct-causation*- (or *DC*-)chain.)

Analysis 1 handles all of the test cases we have considered save the transitivity problem. I now propose a Lewisian solution to that problem.

2.3 Resolving the transitivity problem

We need to appeal to DC-chains in order to tackle test case 1. But there is a DC-chain between the dog-attack and the explosion in our counterexample to transitivity: dog-attack, Patel's detonating of the bomb, explosion. So it must be *only under certain conditions* that the existence of a DC-chain between an event *c* and an event *e* is sufficient for *c*'s causing *e*.

There is an obvious candidate for such a condition. *c** in test case 1 is an early chance-raiser of *e** *in the following special sense*: the chance, assessed at time t_c^+ (immediately after *c**'s actual occurrence), of *e**'s occurring is much higher than it would have been if *c** had not occurred *and none of f*, g* and e* had occurred by* t_e^-. By contrast, the dog-attack in test case 3 is not an early chance-raiser of the explosion: the chance of the explosion's occurring, assessed at time t_D^+ (immediately after the dog-attack) is *not* higher than it would have been if the attack had not occurred and neither Patel's pressing of the button nor the explosion had occurred by t_D^+. So, it is prima facie plausible that causes must be early chance-raisers (in the above sense) of their effects.[6]

The proposal is promising, but, it delivers the wrong verdict in the Figure 9.3 example. Here, *b** causes *e** but it in fact *lowers* the chance of *e**'s occurring. I think Lewis's notion of *quasi-dependence* (1986: 205 ff.) may provide the missing ingredient. In his (attempted) solution to the problem of *late-preemption* under the assumption of determinism Lewis suggests that whether a chain of events occurring in a spatiotemporal region constitutes a causal process or not depends only on the *intrinsic character* of the chain itself, and on the relevant laws. While a course of events in the actual world may not form a chain of dependent events it may still, in its intrinsic character, be just like courses of events in other regions (of this or other worlds with the same laws) that *do* form such a chain. Suppose there is a chain of actual events, *X*, with *c* at the beginning and *e* at the end; *e quasi-depends on c* if the great majority of chains of events intrinsically similar to *X*, as measured by variety of the surroundings, exhibit the proper pattern of dependence (1986: 206).

I propose an analogous notion of *quasi-early chance-raising*. First, a definition of the variety of early chance-raising that concerns us. For any actual events *c* and *e*, *c* is an *early chance-raiser* of *e iff* there is a chain of actual events, $[c, d_1, \dots, d_n, e]$, such that $ch(\text{at } t_e^+, e)$ would have been much smaller than it actually is if *c* had not occurred and none of $d_1, \dots, d_n$ and *e* had occurred by t_c^+. *c* is a *quasi-early chance-raiser* of *e* if there is a chain of actual events, $X = [c, x_1, \dots, x_n, e]$, where the majority of chains of events intrinsically similar to *X*, as measured by the variety of surroundings, and so on, are such that the first event in the chain (the *c*-counterpart) is an early chance-raiser of the final event (the *e*-counterpart).

Now, consider Figure 9.3 again. Although *b** is not an early chance-raiser of *e**, it is a quasi-early chance-raiser (a *quasi-raiser* for short) of *e**. *c** in test case 1 is likewise a quasi-raiser of *e**. But, importantly, the dog-attack in test case 3 is not a quasi-raiser of the explosion.

Here, then, is the analysis that emerges (beginning with a replacement for the definition of direct causation).

*Defn4**: For any actual events *c* and *e*, *c* is a *candidate cause* of *e iff c* is a late and timely chance-raiser of *e*. (Accordingly, we will now speak of *candidate causal chains* (*CC-chains*) rather than direct causation chains.)

> *Analysis 2*
> For any actual events *c* and *e*, *c causes e iff* there is a candidate causal chain linking *c* and *e*, *and c* is a quasi early chance-raiser of *e*.

Analysis 2 handles all the problem cases discussed so far, and it appeals to nothing more than what is already available within Lewis's general theory. In the next section the account is developed further to handle examples of *late* preemptive causation.

3 Late preemption: sigma-tizing the account

The Figure 9.3 example depicts a case of *early* preemption; it is handled by appeal to causal-chains both in Lewis's original account (LPA) and in Analysis 2 above. Cases of *late* preemption, however, have proved much harder to tackle. I will consider two examples. Assume the world is *deterministic* in these two cases.

In the example shown in Figure 9.4, the *b*–*e*-process runs to completion fractionally ahead of the *a*-process, and the firing of *e* prevents *d* from firing. Crucially, if *e* had not fired when it did, *d would have* fired and brought about *e*'s firing a bit later. Neither LPA nor Analysis 2 delivers the correct verdict that *b** is a cause of *e**; the root of the problem is that *g** is not even a candidate cause. For instance, if *g* had not fired (and *e* had not fired by t_e^-), the chance at t_e^-, of *e*'s firing would have been one, and therefore not much smaller than it actually is.

The diagram in Figure 9.5 shows we can have preemption even when the preempted process *does* run to completion (this is preemption *without 'cutting'*, as Lewis (2000) calls it).

The set-up is as before, except for the fact that *e** does not inhibit *d**, and all the events in the *a*–*e*-process occur. But, as before, the *b*-process runs slightly ahead of the *a*-process and *e* fires earlier than it would have done if *d** had been responsible. (Thus,

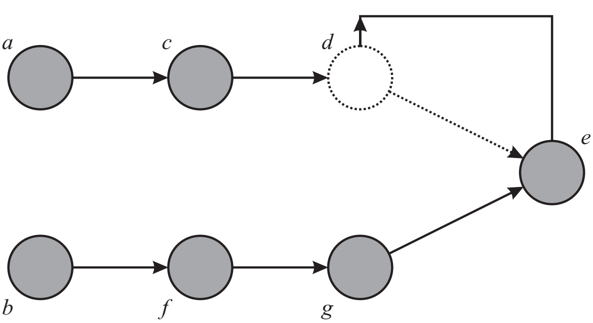

Figure 9.4 Late preemption

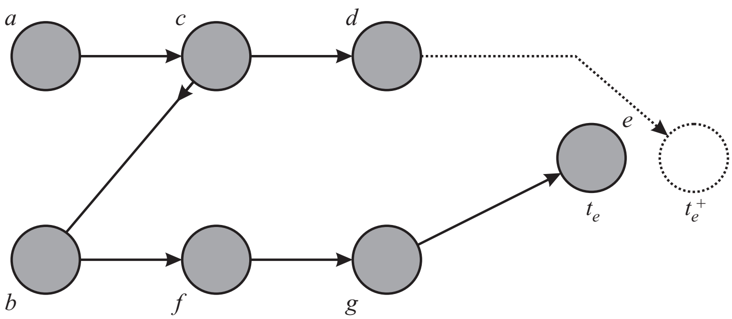

Figure 9.5 Preemption without 'cutting'

the blank circle at t_e^+ depicts the possible later firing of *e*.) Again, *g** is not even a candidate cause *e**: if *g** had not occurred, and *e** had not occurred by t_e^-, the chance of *e*'s firing would have been one, and therefore not smaller than it actually is.

My strategy for dealing with these cases adapts the 'sigma-dependence' approach developed in Ganeri, Noordhof and Ramachandran (1996, 1998) and Noordhof (1999). I will specify the revisions I propose to handle the above problems straight away and explicate them afterwards.

*Defn2**: For any actual events *c* and *e*, and any set of possible events Σ, *c* is a *late* Σ-*chance-raiser* of *e iff* there are values *x* and *y* such that:

(2_1) if *c* were to occur without any of the events in Σ, and *e* were not to occur by t_e^-, then *ch*(at t_e^-, *e*) would be at least *x*;
(2_2) if neither *c* nor any of the events in Σ were to occur, and *e* were not to occur by t_e^-, then *ch*(at t_e^-, *e*) would be at most *y*; and
(2_3) *x*/*y* is large.

Informally, these conditions require that *c would have* been a late chance-raiser of *e but for the possible occurrence of the events in* Σ.

*Defn3**: For any actual events *c* and *e*, and any set of possible events Σ, *c* is a *timely* Σ-*chance-raiser* of *e iff* there are values *x* and *y* such that:

(3_1) if *c* were to occur without any of the events in Σ, then *ch*(at t_e^-, $<e, t_e>$) would be at least *x*;
(3_2) if neither *c* nor any of the events in Σ were to occur, then *ch*(at t_e^-, $<e, t_e>$) would be at most *y*; and
(3_3) *x*/*y* is large.

Informally, these conditions require that *c would have* been a timely chance-raiser of *e but for the possible occurrence of the events in* Σ.

*Defn4***: For any actual events *c* and *e*, *c* is a *candidate cause* of *e iff* there is a set of possible events, Σ, such that:

> (4_1) *c* is a late and timely Σ-chance-raiser of *e*; and
> (4_2) no *non-actual* event is a late Σ-chance-raiser of *e*.

As before, *c causes e iff* there is a candidate causal chain linking *c* to *e* and *c* is a quasi-early chance-raiser of *e*.

Now, let's see how these revised definitions work. In the examples in Figures 9.4 and 9.5, there is a Σ-set such that *b**, *f** and *g** come out as late and timely Σ-chance-raisers of *e**. Consider *b** and take Σ = {*d**}; if neither *b** nor any of the events in Σ (in other words *d**) had occurred, then *ch*(at t_e^-, *e*) and *ch*(at t_e^-, <*e*, t_e>) would have been zero – much smaller than they would have been if *b** were to occur without *d**. By contrast, there is no Σ-set whereby an event in the *preempted* processes (*a**, *c** or *d**) comes out as a *timely* Σ-chance-raiser of *e**. Thus, none of *these* events count as candidate causes of *e**.

The need for condition (4_2) arises in cases such as the Figure 9.3 set-up under the assumption of determinism. In this case, *a**, an event in the *preempted* process, does meet condition (4_1). Take Σ = {*g**}, for instance. If neither *a** nor *g** had occurred, then *ch*(at t_e^-, *e*) and *ch*(at t_e^-, <*e*, t_e>) would have been zero – much smaller than they would have been if *a** had occurred without *g**; so, *a** is a late and timely Σ-chance-raiser of *e**. *However*, there is also a *non-actual* event that is also a late Σ-chance-raiser of *e** for this Σ-set – for example, *c**: *ch*(at t_e^-, *e*) would have been one if *c** had occurred without *g** and would have been zero if neither had occurred. Likewise, *c** will be a late Σ-chance raiser for *e** for *any* Σ-set where *a** is a late Σ-chance-raiser for *e**. So, condition (4_2) blocks *a** from coming out as a candidate cause of *e**, as required.

It remains to check how the analysis delivers the verdict that *b**, say, *is* a cause of *e** in these preemption examples. Let us focus on Figure 9.5. First of all, we need to show that *g** is a candidate cause of *e**. Take Σ = {*d**}; *g** is a late and timely Σ-chance-raiser of *e** – so, condition (4_1) is met. As for (4_2), no *non-actual* event comes out as a late Σ-chance-raiser of *e**. To see why, consider any non-actual event *n*; if neither *n* nor *d** had occurred, *ch*(at t_e^-, *e*) might be one, because the nearest worlds in which the antecedent is true will include worlds in which *g** occurs; thus, *n* will not come out as a Σ-late chance-raiser of *e**. So, *g** is a candidate cause of *e**. Taking Σ = the empty set, ∅, we get *b** and *f** meeting (4_1) too; and (4_2) is trivially met. We thereby have a CC-chain between *b** and *e**, and it is easy to see that *b** is a quasi-early chance-raiser of *e**. Hence, *b** comes out as a cause of *e**, as desired.

Matters are not so straightforward if the assumption of determinism is dropped, however. Suppose the axons in the *b*–*e*-chain are very unreliable whereas the ones in the *a*–*e*-chain are reliable. Now, for all that has been said, there might be a non-actual event *n*, independent of the *a*- and *b*- processes, which, had it occurred, would *definitely* have brought *e** about. In such a case, *g** will not be a late Σ-chance-raiser of *e** where Σ = {*d**}. So, one might wonder how *g** can come out as a candidate cause of e^* in this scenario. The solution is straightforward: we simply put n along with d^* (and any other potential chance-raisers of e^*) into our Σ-set. g^* *would be* a late and timely chance-raiser of e^* but for the occurrence of the events in this enlarged Σ-set, Σ^*; that is, g^* will be a late and timely Σ^*-chance-raiser of e^*. The same strategy suffices to establish a CC-chain between b^* and e^* and to get b^* coming out as a cause of e^*.

The account we have arrived at handles most varieties of 'redundant' causation – causation involving over-determination or preemption. There remain cases of 'trumping' preemption (see Schaffer 2000a) to consider, and the account needs further modification to handle cases of *inhibition-of-inhibitors*.[7] But this account achieves more than many have thought possible within Lewis's original framework. Lewis (2000) himself, for example, feels compelled to take 'when-when' and 'how-how' dependence as sufficient for causation – thereby, rendering mere hasteners and delayers causes. And causation is still transitive on his new account! I leave it to readers to compare my account with other counterfactual theories – such as Noordhof (1999), Lewis (2000), Schaffer (2001a) and Barker (this volume). There are interesting differences between the accounts, but I contend the account presented here stands up to comparison.[8]

Notes

1 See, for example, my earlier self (Ramachandran 1997) and Barker (this volume) for attempts that eschew the chance-raising line.
2 See for example Menzies (1989; 1996), Paul (1998a) and Noordhof (1999).
3 This simply follows from the fact that when the chance of an event, e, is assessed at a time, t, one takes the history of the world up to t into account. So, the chance *at time t* of e's occurring is effectively the *conditional* chance of e's occurring *given the history of the world prior to t*.
4 This line is adopted in Ramachandran (1998) and Noordhof (1999).
5 If e were to occur before t_e^-, $ch(\text{at } t_e^-, <e, t_e>)$ would be zero, not one. So there is no need to restrict our attention to worlds in which c does not occur and e does not occur by t_e^-.
6 This proposal should be distinguished from Beebee's (this volume) requirement that causes be (early) chance-raisers of their effects. For she is talking of early chance-raisers in *Lewis's* sense.
7 Barker (this volume) appears to have a solution to the trumping preemption problem that can be adopted by sigma-dependence approaches such as ours – see Noordhof (this volume). As for the inhibition problem, inhibitors of inhibitors of an event e should, to my mind, count as causes of e; but they do not come out as such by the present account. I believe a straightforward modification to the definition of 'candidate cause' can deal with this problem. But the issue calls for a wider discussion beyond the scope of this chapter.
8 Thanks to Matt Densley, Chris Foulds, Carl Hoefer, Simon Langford, Kyle Murray, Paul Noordhof, Sonny Ramachandran, and Ken Turner for comments on earlier drafts.

Chapter 10

Probabilistic cause, edge conditions, late preemption and discrete cases

Igal Kvart

Key words: ab initio probability increase; causal relevance; cause; chance; decreaser; electron trajectory; exhaustive channellers; increaser; neutralizer; overlapping; preemption; probabilistic relevance; radioactive decay; screener; stable increaser; stable screener.

In this chapter I attempt to resolve a number of issues concerning the analysis of token cause, and in particular of token causal relevance, from the perspective of the analyses of these notions that I proposed elsewhere.[1] In the first part of the chapter (Sections 1–4) I summarize the probabilistic analyses of token cause and of causal relevance. The probabilities employed are chances. The analysis of cause (Section 1) focuses on specifying the right notion of probability increase. But causal relevance, I argue, is a crucial prerequisite for being a cause and accordingly a crucial ingredient of the notion of cause (and of counterfactuals), and is central to a correct analysis of preemption. The main idea in the analysis of causal relevance (Section 2–4) is that, in a chancy world, causal irrelevance is secured either through probabilistic irrelevance or through the presence of a so-called causal-relevance *neutralizer*. Essentially, as I explain below, a causal-relevance neutralizer screens off A from C in a stable way and is such that A is not a cause of it. Yet, despite appearances to the contrary, the account is not circular.

I then proceed to discuss the following issues concerning this conception regarding one token event A being a cause of a later token event C. In Section 5 I argue for a constraint regarding neutralizers pertaining all the way to the upper edge of the occurrence time of the C-event. The strong version of the problem of late preemption is analysed in Section 6. One general outcome of the analysis of causal relevance presented here is that in preemption cases (early or late), the preempted cause is not a cause since it is causally irrelevant to the effect, and that neutralizers are often succinctly specifiable. Causes in what might be taken to be cases with discrete time, such as cases of radioactive decay or of electrons moving between different trajectories, pose challenges for a probabilistic analysis such as the above regarding suitable intermediate events. This is so since it may seem that either there is no room at all for intermediate events of the requisite sort, or else that the only available intermediate events appear to be simultaneous with the effect. I discuss two such challenges in Sections 7–8.[2]

1 Cause

I assume an indeterministic, chancy world, where the chance of *C* is conditional on some prior world-state or history (and possibly other events). I thus assume a chance function $P(C/W_t)$, where W_t is the world history up to *t* (or, if you will, a world-state just prior to *t*), and *t* is earlier than the beginning of t_C, which is the interval to which *C* pertains. I have argued that for *A* to be a token cause of *C*, *A* must raise the chance of *C ex post facto*, that is, while taking into account not only the world history prior to *A*, symbolized as W_A, but also the world history during the interval between *A* and *C*.[3] The use of probabilities in this paper is always in this sense of chance, a point that is important to bear in mind. Traditionally, probabilistic analyses of causation had to face the difficulty of how to reconcile the intuitive idea that causes raise the probability of their effects with cases where causes seem to lower the probability of their effects.[4] On the token level, the natural way of expressing probability increase in terms of chances is:

(1) $P(C/A.W_A) > P(C/{\sim}A.W_A)$

(1) is called *ab initio probability increase* (in short: *aipi*). For instance, suppose I dropped the chalk (*A*), and the chalk then fell on the floor (*C*) (absent any complications). *A*, intuitively, was a cause of *C*, and ab initio probability increase obtains.

However, in (1) only the history of the world prior to *A* is taken into account. Ex post facto probability increase, on the other hand, must take into account the intermediate history as well – the history pertaining to the interval between *A* and *C*. The condition for being a cause, then, is not ab initio probability increase, which is *not* sufficient for being a cause,[5] but rather a condition of ex post facto probability increase, and in particular a specific form of it, which we shall spell out now. (Note that W_A is the world history, considered from a realist perspective, not to be conflated with the epistemic issue of how much of it any particular individual knows.) Thus, consider Example 1.

Example 1

The Comeback Team had been weak for quite a while, with poor chances of improvement during the following season. Consequently, there were very high odds of its not coming out first. Nevertheless, *x* bet \$*Y* on its coming out first (*A*). However, later, but before the beginning of the games, the Comeback Team was acquired by a new wealthy owner, an event which had been quite unlikely at the time of the bet. The new owner subsequently also acquired a few first-rate players. Consequently, the team's performance was the best in the season (*E*), *x* won her bet, and *C* occurred: *x* improved her financial position.[6]

As of t_A, *A* yielded a probability decrease of *C*, that is, ab initio probability decrease. But given *E*, *A* yielded a higher chance of *C* (see (2) on p. 165). And indeed, intuitively, *A* was surely a cause of *C*.

Let us spell out this notion of *A*'s yielding a higher chance of *C* ex post facto, and at the same time illustrate ex post facto probability increase in cases of *ab initio probability decrease* (the latter being (1) with '<' instead of '>'; in short, *aipd*). The main idea of ex post facto probability increase despite ab initio probability decrease that is suitable for the notion of cause is that there must be an actual intermediate event E that, when held fixed – that is, added to the conditions in both sides of the ab initio probability decrease condition – yields probability increase. In other words:

(2) $P(C/A.E.W_A) > P(C/{\sim}A.E.W_A)$

And indeed, E in Example 1 above satisfies condition (2). Call an event such as E an *increaser*. In Example 1, E is an increaser for A and C.

A note of caution: the term 'increaser' should not lead you to think that an increaser increases the chance of C when added to the condition in $P(C/A.W_A)$. The import of an increaser E as such does *not* involve a characterization of the relation between the above conditional probability vs. the above conditional probability with E added to the condition. Rather, what is at stake is the relation of two conditional probabilities, *both with E* in the condition, one with A, the other with ${\sim}A$. For instance, consider a student x who took an exam:

A – x gave a wrong answer to question b.

Yet:

E – x answered question d correctly.
C – x received a high grade.[7]

A yields ab initio probability decrease for C. Yet E is not an increaser since:

$P(C/A.E.W_A) < P(C/{\sim}A.E.W_A)$

So E does *not* yield a higher probability of C when held fixed in the condition on both sides of the ab initio probability condition. Yet E yields a higher probability of C given A (and W_A) when it is *not* held fixed. In other words:

$P(C/A.E.W_A) > P(C/A.W_A)$

Thus, E is not an increaser (for A and C).

However, ab initio probability increase for A and C need not yield that A is a cause of C, since there might be a decreaser for A and C (that is, an intermediate E fulfilling (2) with '<' instead of '>'), thereby undermining the indication of ex post facto probability increase by the ab initio probability increase. Hence ab initio probability increase is not sufficient for being a cause. That raising the probability is not a sufficient condition for being a cause, even without the cutting of causal routes, has not been appropriately heeded.[8]

The same problem may plague the presence of an increaser: an increaser E might have a further decreaser (for it), that is, an intermediate event[9] F fulfilling:

$$(3)\ P(C/A.E.F.W_A) < P\ (C/{\sim}A.E.F.W_A)$$

(3) undermines the indication of ex post facto probability increase yielded by the increaser E. The possibility of there being a decreaser for a given increaser shows that the indication of ex post facto probability increase yielded by an increaser need not be sustained when other intermediate events are also taken into account. In order for A to be a cause of C it must have an increaser *E without* a further decreaser for it (such as F in (3)). Call such an increaser *stable* (or *strict*).[10] The probability increase indicated by a stable increaser is indeed stable since it is not reversed when other intermediate events are taken into account.

It is plausible to expect the notion of cause to consist primarily of probability increase that is both ex post facto and stable; and these requirements are indeed extensionally adequate. The notion of A's raising the probability of C needed for the truth conditions of A's being a cause of C must be sufficiently resilient so that the feature of probability increase involved proves stable across the intermediate history. Given a realist position regarding facts and chances,[11] a fully fledged epistemic correlate of A's being a cause of C is its assessment as such from a retrospective perspective based on information concerning the facts and the chances as of some time after the end of t_C. The epistemic correlate of ab initio probability increase, when relativized to information possessed by a cognizer of facts that pertain up to t_A (in contrast with the non-epistemic W_A) and to the cognizer's subjective probability conceived as assessment of chances as of t_A is useful for prediction when the cognizer is at t_A and assesses whether A will be a cause of C: this is an ab initio *assessment*.

For the sake of a uniform terminology, consider the case of ab initio probability increase as a case of a null increaser (as in the example above of the chalk that fell to the floor), and similarly consider a case of ab initio probability decrease as a case of a *null* decreaser. A null increaser may be stable (that is, if there is no decreaser). In such a case, its presence constitutes a sufficient condition for being a cause.[12]

2 Causal-relevance neutralizers

The existence of a stable increaser is a necessary, but not sufficient, condition for being a cause. Causal relevance (of A to C) is also a necessary condition for A's being a cause of C. The presence of a stable increaser for A and C yields A as a *prima facie* cause of C, and renders A a cause of *C simpliciter* so long as A is causally relevant to C. Thus, if A is not causally relevant to C, A is not a cause of C even if there is a stable increaser for A and C. A may be causally irrelevant to C despite the presence of a stable increaser. This would be the case if there is an intermediate event that *neutralizes* the would-be causal relevance of A to C. The presence of a stable increaser therefore suffices neither for causal relevance nor for being a cause.

Consider Example 2.

Example 2

x was pursued by two enemies who wanted to kill him. They discovered that he was going to be at a particular time at a particular meeting place in an area covered with heavy snow. Enemy 1 arrived with his attack dog, discovered that there was a cave with a very small entrance close to the meeting place, and hid there with his dog. Enemy 2, with his gun ready, found another place to hide overlooking the meeting place. At the designated time, x arrived at the meeting place, and indeed:

A – Enemy 1 released his dog at t_1.

However, due to the mass of snow that covered the slope above the entrance of the cave, E occurred:

E – an avalanche completely blocked the entrance of the cave at $t_1 + dt$.

E occurred in the nick of time, before the dog had a chance to rush forward,[13] and so both Enemy 1 and his dog were trapped in the blocked cave. However, when Enemy 2 observed the arrival of x, B occurred:

B – Enemy 2 shot at x.

Enemy 2 indeed hit x, and consequently C occurred:

C – x was injured.

Intuitively, A was not a cause of C, but B was. However, A bore ab initio probability increase to C, that is:

$$\mathrm{P}(C/A.W_A) > \mathrm{P}(C/{\sim}A.W_A)$$

with no decreaser, so A had a null stable increaser vis-à-vis C. But A was nonetheless not a cause of C, since, intuitively, A ended up being causally irrelevant to C.[14]

In general, for A to be causally relevant to C, A must be *probabilistically relevant* to C. A strong conviction to this effect lies at the heart of a full-blooded probabilistic approach to causal relevance. A is *probabilistically relevant* to C just in case there is either an increaser or a decreaser for A and C. Increasers and decreasers may, as noted, be null or not. There is either a null increaser or a null decreaser for A and C if and only if:

(4) $P(C/A.W_A) \neq P(C/{\sim}A.W_A)$

A non-null increaser or decreaser F for A and C yields:

(5) $P(C/A.F.W_A) \neq P(C/{\sim}A.F.W_A)$

Probabilistic relevance yields a *prima facie* case of causal relevance (which is overruled if there is a causal-relevance neutralizer – see p. 170).

When the ab initio inequality (4) does not hold but (5) does, call an intermediate event F fulfilling (5) a *differentiator* (for A and C). When the ab initio inequality (4) holds, consider an empty intermediate event a *null* differentiator (for A and C). Thus, a null differentiator is either a null increaser or a null decreaser. The absence of a null differentiator amounts to (4) not holding, in other words, it amounts to the presence of probabilistic equality; that is:

Equi-Probability: $P(C/A.W_A) = P(C/{\sim}A.W_A)$

A differentiator (null or not) is either an increaser or a decreaser (null or not). Accordingly, A is probabilistically relevant to C just in case A makes a probabilistic difference to C, either directly (as in (4)), or via an intermediate event that is held fixed (as in (5)). A is therefore probabilistically relevant to C just in case there is a differentiator (null or not) for A and C.

Without probabilistic relevance there is no causal relevance, and yet A was probabilistically relevant to C in the last example (Example 2). What accounts, then, for the causal irrelevance of A to C there? It is, I propose, the presence of an intermediate event that attests to the would-be causal relevance of A to C being *neutralized*. In that example E was such an intermediate event. Call such an event a *causal-relevance neutralizer* (in short: a *neutralizer*, or a *crn*). So if A is probabilistically relevant to C, A is causally relevant to C just in case there is no causal-relevance neutralizer (for A and C). Our task now, then, is to characterize what renders an intermediate event a causal-relevance neutralizer.

A would-be causal chain from A to C is neutralized if it is cut off; but the latter is not the only way for it to be neutralized. Such a chain may be diverted, or it may simply dissipate. Yet note that a causal-relevance neutralizer need not be the event that cuts off would-be causal chains from A to C even when such chains are cut off. The role of a causal-relevance neutralizer is to secure that no would-be causal chains[15] from A proceed all the way to C, thereby securing the absence of causal relevance.

3 Candidates for causal-relevance neutralizers

A causal-relevance neutralizer E, which secures causal irrelevance of A to C (despite probabilistic relevance), must, I suggest, *screen off A from C*; that is, it must fulfil:

(6) $P(C/A.E.W_A) = P(C/{\sim}A.E.W_A)$

An event *E* satisfying (6) is a *screener* for *A* and *C*.[16]

As an illustration, consider Example 3.

Example 3

Consider a pipe that leads from a main tap (faucet) to a pool. This main tap can be in only one of two positions: in one position it allows water to flow freely through the pipe, whereas in the other position it doesn't allow any water to flow through the pipe. Furthermore, the pipe has at an intermediate point its own separate tap that also controls the water flow through it. This intermediate tap also has two positions, of the same sort as in the main tap. Assume that it was in the open position. The main tap was originally closed. But then *x* switched the main tap to the open position (*A*). And indeed, the pool filled up (*C*). Assume that the chance, as of t_A, of further interventions regarding the positions of the taps (other than *A*) is small. Hence there is ab initio probability increase for *A* and *C* (and thus there is a null differentiator for them), and *A* is therefore probabilistically relevant to *C*. Assuming that there are no other pertinent unexpected aspects of the story took place, *A* is also intuitively causally relevant to *C* as well as a cause of *C*.

Now let us move to a variation in which, unlike the previous version, immediately after *A*, in an unrelated development, *E* occurred: the intermediate tap was closed. Consequently, no water originating from the main tap reached the pool. Yet, it started to rain around that time, and the rain filled up the pool. We can assume further that, as of t_A, and given *E*, the chance of *C* given *A* vs. ~*A* is the same. Assume also that no other pertinent unexpected occurrences took place. Intuitively, then, *A* was not a cause of *C*, and indeed, as it turned out, in view of *E*, *A* ended up being causally irrelevant to *C*. *E* is indeed a screener for *A* and *C*, and, I suggest, under circumstances akin to the rough specification of the above sort, *E* is also a main component of a causal-relevance neutralizer for *A* and *C* (see *E′* on p. 170).

Yet the mere presence of a screener *E* need not yield a *stable* ex post facto probabilistic equality, since a screener may have a differentiator *for it*. But, I suggest, a causal-relevance neutralizer for *A* and *C* must also screen off *A* from *C* in a stable, that is, unreversed, way. That is, there must not be any other intermediate event *F* that undoes this screening-off. In other words, there must not be an intermediate event *F* such that:

(7) $P(C/A.E.F.W_A) \neq P(C/{\sim}A.E.F.W_A)$

If *E* is a screener for *A* and *C*, fulfilling (6), for which there is no intermediate *F*, fulfilling (7), then the probabilistic equality of (6) is indeed stable. Call such a screener *E* (satisfying (6)) for which there is no such *F* satisfying (7) a *stable screener* for *A* and *C*. A stable screener, then, yields stable probabilistic equality.[17]

To illustrate, let us move now to a still further variation of Example 3 above, in which, in addition to the story of the last variation, the following *F* occurred: the intermediate tap was re-opened (at time *t*). *t* was later than *A* and *E* but well before *C*. (Recall: *A* was: *x* switched the main tap to the open position; *E* was: the intermediate tap was closed.) Given that the intermediate tap was re-opened, there would be water flow from it to the pool, given *A*, and *A* thus intuitively would end up being causally relevant to and a cause of the actual *C* (which was: the pool filled up).[18] In this variation, *E*, despite being a screener for *A* and *C*, was not a stable screener, since *F* is a differentiator for *E* (vis-à-vis *A* and *C*). A screener that is not stable falls short of securing causal irrelevance.

If, however, as in the previous version, neither *F* nor any other differentiator for *E* occurred, then *E* is a stable screener for *A* and *C*. As noted, in such a case, under the circumstances, *A* would end up being causally irrelevant to *C*, and *E* is, as I suggested, a main component of a causal-relevance neutralizer for *A* and *C*. This requirement, that a candidate for a causal-relevance neutralizer for *A* and *C* be a stable screener for them, reflects a probabilistic approach to causal relevance, viewed as a probabilistic phenomenon. And indeed, an extension *E′* of *E* is a stable screener for *A* and *C* and also a causal-relevance neutralizer for them, where *E′* is: the intermediate tap was closed and remained so.[19]

Consider again Example 2 in Section 2 in which, intuitively, *A* (Enemy 1 released his dog at t_1) ended up being causally irrelevant to *C* (*x* was injured). Yet there is a null differentiator for *A* and *C*, and thus *A* was probabilistically relevant to *C*. But *E* (an avalanche completely blocked the entrance to the cave at $t_1 + dt$) was a stable screener for *A* and *C*: given *E*, it makes no difference to the probability of *C* whether *A* or ~*A* took place, and there is no differentiator for *E* (vis-à-vis *A* and *C*). And indeed, *E* is a causal-relevance neutralizer in this case.

If *E* is a stable screener (for *A* and *C*), so would be true conjunctive expansions of *E*. In order to focus on the intrinsic features of causal-relevance neutralizers, we shall confine our attention to *lean* stable screeners, in other words, stable screeners devoid of extra information that plays no role in their being stable screeners.[20] Call lean stable screeners candidates for *causal-relevance neutralizer*s (or for short: *crn-candidates*).

4 The analysis of causal-relevance neutralizers

What then are causal-relevance neutralizers? As noted, in accordance with the spirit of our probabilistic approach, a causal-relevance neutralizer *E* for *A* and *C* must be a stable screener for them. But stable screeners are abundant.[21] Yet local stable screeners can play an important role other than that of causal-relevance neutralizers. Assume a straightforward case of only one causal route between *A* and *C*, where *A* is a cause of *C*. Thus, consider Example 4.

Example 4

> $A - x$ fired at y at t_0.
>
> $C - y$ was hit at t.

Assume no complications: x's bullet hit y. Consider the trajectory of the bullet, and select an arbitrary intermediate temporal point t_1 along the bullet's trajectory. Consider event E_1:

> $E_1 - x$'s bullet was at t_1 at point p_1 in mid-air, with momentum m_1, spin s_1, and so on.[22]

E_1 is a stable screener for A and C. But of course E_1 is not a causal-relevance neutralizer for A and C: A is surely a cause of C and thus causally relevant to it. Rather, E_1 *exhaustively channels* the causal relevance of A to C.[23]

So stable screeners can be both causal-relevance neutralizers and exhaustive channellers. Yet what is typical of an exhaustive channeller E (for A and C) is that A is a cause of E. To set apart, then, causal-relevance neutralizers from exhaustive channellers, I propose the following thesis:

> Thesis: E is a causal-relevance neutralizer for A and C just in case E is a lean stable screener for A and C of which A is not a cause.

This thesis threatens circularity, since the notion of cause is employed, but this threat of circularity will evaporate, as we will see below. So probabilistic relevance yields a prima facie case of causal relevance, which is undercut if and only if there is a causal-relevance neutralizer.

Let us now apply the above analysis to Example 2 above, with the dog and the avalanche. Intuitively, A (Enemy 1 released his dog at t_1) ended up being causally irrelevant to C (x was injured), even though there is a null differentiator for A and C. And indeed, E (an avalanche completely blocked the entrance to the cave at $t_1 + dt$) was a stable screener for A and C, as noted above, and intuitively A was not a cause of E (E occurred independently of A). Indeed, in terms of our analysis, there was also no differentiator for A and E, and thus A was probabilistically irrelevant to E and therefore also causally irrelevant to E. Hence E is a causal-relevance neutralizer for A and C. We shall henceforth call a causal-relevance neutralizer for short just a *neutralizer*. Yet B (Enemy 2 shot at x) has a null stable increaser for C, and hence there is a differentiator for B and C. Further, there is no neutralizer for B and C. B, therefore, intuitively as well as by the above analysis, ended up being a cause of C, whereas A was not, since A ended up being causally irrelevant to C.[24]

In order to illustrate the above condition and address the issue of the risk of infinite regress, consider Example 4 on p. 170 in which A was: x fired at y at t_0, and C was: y was hit at t. In this case (which is, again, a straightforward case with no complications), A was intuitively a cause of C, and hence A was causally relevant to C. Surely there is ab initio probability increase of A to C, and hence there is a null differentiator. So in checking for causal relevance of A to C, in terms of our analysis, we must look for a candidate for a neutralizer for A and C. And indeed, E_1 above (x's bullet was at t_1 at point p_1 in mid-air, with momentum m_1, spin s_1, and so on) screens off A from C, and is in fact a stable screener for them, and is thus a candidate for a neutralizer. But, intuitively, A is surely a cause of E_1. And indeed, there is ab initio probability increase for A and E_1, and accordingly there is a null stable increaser for them. So E_1 is not a neutralizer for A and C unless A is causally irrelevant to E_1, that is, unless there is a neutralizer E_2 for A and E_1.

So consider E_2, which is just like E_1 only with t_2, p_2, m_2, s_2, and so on instead of t_1, p_1, m_1, s_1, where t_2 is some intermediate point after A and before t_1. E_2 is a candidate for a neutralizer for A and E_1 just as E_1 was a candidate for a neutralizer for A and C. But again, A is intuitively a cause of E_2, and indeed, A has a null stable increaser to E_2. Hence E_2 is not a neutralizer for A and E_1 unless there is a neutralizer E_3 for A and E_2; and so on and so forth.

What happens if the chain does not terminate? And more generally, what happens if no such chain terminates? In such a case, there is no neutralizer for A and C. Recall now that this is a prima facie case of causal relevance of A to C since there is a differentiator for them and probabilistic relevance yields a prima facie case of causal relevance. But since there is no neutralizer to overrule the prima facie causal relevance of A to C, A is therefore causally relevant to C. Thus, there may be infinite regress, but there is no circularity, since the infinite regress case is a case of causal relevance.

This sort of pattern holds in general. If such a chain of intermediate candidates for neutralizers E_is terminates, each E_i is a neutralizer of A and E_{i-1}, and hence E_1 is a neutralizer for A and C, and thus A is causally irrelevant to C. Otherwise, there is an infinite series of such E_is[25] where the corresponding decreasing t_is converge to some temporal point between A and C. If no such chain terminates, if there is only infinite regress of this sort, then A is causally relevant to C since in such a case there is no neutralizer for A and C (and yet there is a differentiator for them). In fact, such infinite regress is a hallmark of causal relevance. Cases with terminated chains have neutralizers, and are thus cases of causal irrelevance. This pattern can be shown, along such lines, to hold in general.[26]

The above presentation of the analyses of cause and of causal relevance summarizes more detailed analyses that are presented elsewhere with more elaborate arguments for their adequacy.[27] The reader who needs to be convinced further by the adequacy of these analyses is advised to look at the more detailed presentations.

5 Neutralizers pertaining up to the upper end of t_C

Neutralizers are intermediate events fulfilling the conditions spelled out above. However, so far we haven't been very specific about what being intermediate precisely means regarding neutralizers vis-à-vis the temporal edges, namely the upper and lower ends of the temporal interval ($t_A t_C$), which is the minimal interval that includes t_A and t_C (which are the time intervals to which A and C pertain). Surely the occurrence time of intermediate events must not start earlier than the lower end of the occurrence time of the *A*-event and must not end later than the upper end of the occurrence time of the *C*-event. But can the occurrence time of intermediate events pertain all the way up to the upper end of the occurrence time of the *C*-event?

There are strong reasons in favour of admitting neutralizers whose occurrence time pertains all the way up to the upper end of the occurrence time of the *C*–event. This issue will be illustrated below in discrete-time-like cases such as radioactive decay (Section 7). However, once convincing considerations in favour of this option are established, it cannot be left unconstrained. The reason is simple: *C* is always a stable screener for *A* and *C*, since $P(C/A.C.W_A) = P(C/{\sim}A.C.W_A) = 1$, and yet *C* itself must not in general qualify as a candidate for being a neutralizer for *A* and *C*. This is so since, if *C* is allowed to qualify as a candidate for being a neutralizer for *A* and *C* without further constraint, then, in particular, in all cases in which *A* is not a cause of *C* for any such pertinent *A*, *A* would come out as causally irrelevant to *C*. But this implies that all prior events that are not causes of *C* are also causally irrelevant to *C*. Yet this is absurd: events that are not causes of a later event may still be causally relevant to it, if in particular they are purely negatively causally relevant to it. And the relation of being purely negatively causally relevant is prevalent.[28]

So allowing neutralizers to extend all the way up to the upper end of the occurrence time of *C* in an unconstrained way would undermine an account of cause of the sort presented above, since it would imply that all cases of purely negative causal relevance are cases of causal irrelevance, and this is absurd. I have argued elsewere[29] that having some positive causal relevance to *C* is tantamount to being a cause of *C*, and of course causal irrelevance rules out being a cause. The remaining group of events earlier than *C*, other than those that bear some positive causal relevance to *C* or are causally irrelevant to *C*, consists of those that do not have any positive causal relevance to the later event *C* and yet are causally relevant to *C*: these are the events that have purely negative causal relevance to *C*.

In a more formal way, causal relevance has been probabilistically characterized above in Sections 2–4 as probabilistic relevance without a neutralizer. Given causal relevance, having some positive causal relevance amounts to there being a stable increaser, and having some negative causal relevance amounts analogously to there being a stable decreaser. Purely negative causal relevance thus amounts to causal relevance (that is, to probabilistic relevance without a neutralizer) with a stable decreaser and without a stable increaser. Consider first the following example of mixed causal relevance:[30]

Example 5

A particular patient was in a very poor shape, suffering from a liver problem as well as a lung problem. He was given a medication (*A*) to help with his liver problem, but the medication had deleterious side effects vis-à-vis his lung condition. Yet the patient's overall health was in an improved condition a while later (C). Thus, A had mixed causal impact on C: it had positive causal relevance in improving the patient's liver condition, but also had negative causal relevance to C in aggravating the patient's lung condition.

Consider now the following straightforward example of purely negative causal relevance.

Example 6

x and y were pulling a rope in opposite directions. x won. But the fact that y pulled the rope in his direction was purely negatively causally relevant to the fact that the rope ended up on x's side.

Yet such events that bear purely negative causal relevance to C would not be recognized as such – they would wrongly be counted as causally irrelevant to C – if we allow C to qualify as a candidate for a neutralizer for itself. And yet surely being purely negatively causally relevant to an event implies being causally relevant to it. Thus, one must impose a constraint on being a candidate for a neutralizer that yields that C must not qualify as a neutralizer for itself.

But this is not enough, since in general this sort of trivialization holds not only for C itself, when considered as a candidate for a neutralizer for itself and A, but also for many events C^* that imply C, or, more precisely, that yield C with probability 1 (given W_A), when considered as candidates for neutralizers for C. Such a C^* that yields C with probability 1 (given W_A) is too a stable screener for A and C. If we then confine the restriction on neutralizers, the occurrence time of which pertains all the way to the upper end of the occurrence time of C in such a way that it excludes only C itself, then, for any A and C, consider an informational expansion of C. An informational expansion of C would be, for instance, a conjunctive expansion of the form $B.C$, where B is an (actual) intermediate event (for A and C). Select B so that A is not a cause of B. Then, for the kind of case under discussion here, A is also not a cause of B. C – A is not a cause of C since A is purely negatively causally relevant to C. If we impose a restriction on candidates for neutralizers so as to exclude only C itself, then such B. C would qualify as a neutralizer for A and C since B. C screens off A from C in a stable way. So the account faces trivialization since, again, if A is not a cause of C, A would always come out as causally irrelevant to C.

Hence an appropriate constraint on neutralizers, when allowing the occurrence time of neutralizers to stretch all the way up to the upper end of the occurrence time of C, is:

(8) A neutralizer E for A and C must fulfil: $P(C/E.W_A) < 1$

Condition (8) rules out C as well as any C^* that yields C with probability 1 (given W_A) as neutralizers for A and C. Making sure that C does not qualify as a neutralizer for A and C is crucial when A has purely negative causal relevance to C, but is harmless or not objectionable in the other cases, namely the cases in which A is causally irrelevant to C or has some positive causal relevance to C. If A has some positive causal relevance to C, A is a cause of C, and thus C would not qualify as a neutralizer for A and C on our analysis of causal relevance without any further constraints. If A is causally irrelevant to C, then, if we see to it that C does not qualify as a neutralizer for A and C, we can expect that there is some intermediate event other than C that serves as a neutralizer for A and C. This is since in cases of causal irrelevance that are not cases of probabilistic irrelevance, the motivation for expecting the presence of a neutralizer anyhow hinges on the expectation that there are intermediate events other than A or C that attest to the would-be causal relevance of A to C being neutralized; and if there is one such neutralizer, then often there is more than one.

Elsewhere[31] I argued that a corresponding constraint must hold also vis-à-vis the possibility that the occurrence time of a neutralizer extends all the way down to the lower end of the occurrence time of A, namely, a constraint to the effect that for E to be a neutralizer for A and C, it must be the case that $P(A/E.W_A) < 1$. (I have argued that an even stronger constraint is well motivated, to the effect that the occurrence time of neutralizers must not stretch all the way down to the lower end of the occurrence time of A. I will not, however, elaborate on this issue here.)

However, we need to be more precise about the relation between the temporal edges of C and of a neutralizer-candidate E. The time to which C pertains must include, but need not coincide with, the occurrence time of the C-event. Thus, the time to which C pertains, its *pertinence* time, in other words t_C may extend beyond the actual occurrence time of the C-event. (This is so since the sentence C may have a temporal quantifier such as ‘during T’ or even, more emphatically, ‘some time during T’.) Since we proceed here in terms of narrow individuation of events, we may replace the terminology of the time to which the sentence that determines the event pertains with the notion of the *specified* occurrence time of the narrowly individuated event in question. These two notions come down to the same thing, and must be distinguished from the notion of the *actual* occurrence time of the event in question. That is, if the C-event in question is not temporally fragile (and is actual – here we are concerned only with actual events), we may contrast its actual occurrence time, the interval *throughout* which it in fact occurred or took place, with its specified occurrence time – the temporal component of the C-event. (An event can be specified as temporally fragile by a phrase such as ‘exactly at t’ that specifies its occurrence time.) One may thus distinguish between temporally fragile events, where the specified occurrence time coincides with the actual occurrence time, and cases where C is temporally non-*fragile*, so that the actual occurrence time of C does not overlap with, but is included in, its specified occurrence time.[32]

The difference between fragile and non-fragile events plays a significant role in counterfactual analyses of cause such as Lewis’s, and in particular in the case of late preemption (see Section 6 on p. 176), since fragile events generally bypass a major problem faced by Lewis’s (1973b) and (1986) accounts of causation; and yet restricting ourselves exclusively to temporally fragile events is an unreasonable limitation.

It is clear that a neutralizer E must be an intermediate event at least in the sense that the upper end of its actual occurrence time must not exceed the upper end of the actual occurrence time of C. This is an admissibility condition for an event to qualify as a neutralizer. And there is a good motivation for allowing the *specified* occurrence time of a neutralizer to extend all the way up to the upper end of the specified occurrence time of C, as we shall see in the next section. Yet another related admissibility condition for being intermediate is that the specified occurrence time of the event in question (the pertinence time of the sentence that determines it) does not exceed the upper end of the specified occurrence time of C. (Similar constraints apply to the lower end of the occurrence time of an intermediate event and the lower end of the occurrence time of A – both specified and actual, respectively. Of course, only actual events have actual occurrence times.) Yet the argument presented above to the effect that a neutralizer E must adhere to the constraint presented in (8) applies in general in cases of coincidence of the temporal upper ends of the occurrence times of E and C, whether actual or specified.

To sum up, the actual occurrence time of a neutralizer E must not exceed that of C, and likewise for specified occurrence times. When the upper ends of the occurrence times of both coincide, whether actual or specified, condition (8) seems to cover the constraint we need to observe.

6 Late preemption again

Example 7

On a standard form of the problem of late preemption, Suzy and Billy threw rocks at a window which shattered, where:

A – Suzy threw her rock during T_1,

B – Billy threw his rock during T_2.

Suzy's rock hit the window, and Billy's passed through right after Suzy's. Elsewhere[33] I applied the above analysis to the problem of late preemption in a version where the effect was: the window shattered at t_3. The T_i's are temporal intervals that may extend, beyond the actual occurrence times. t_3 is the exact occurrence time of the specified event. However, some challenged that this was the easier problem. The tougher problem of late preemption is when the effect is temporally not fragile.[34] Indeed, for counterfactual theories of causation the latter is the real challenge. In an analysis employing narrow individuation applied to a non-fragile version, as here, the effect should therefore be taken to occur *within* a specified interval that may extend beyond its actual occurrence time, as T_3 below. So consider:

C – the window shattered during T_3.

Whereas A was a cause of C, B was not: Suzy's rock hit the window, Billy's did not. Now consider:

H – Billy's rock didn't hit the window until t_{C^-}).

t_{C^-} is the upper end of t_C. Thus, $t_{H^-} = t_{C^-}$. (We can replace 'until t_{C^-}' in H by: until the upper part of T_3.)

H screens B off from C, and stably, and B lowers the chance of H, with no stable increaser, and thus B is not a cause of H. So H is a neutralizer for B and C, and indeed constraint (8) presented in Section 5 is satisfied: the probability of C given H and W_B is less than 1. Hence B is causally irrelevant to C and thus is not a cause of C,[35] and this of course, intuitively, is the right outcome.[36]

Thus, our analysis offers a treatment of the version of late preemption with a temporally non-fragile effect (as well as with non-fragile cause-candidates) and with the neutralizer H above which is content-wise most natural.[37] (The same treatment also applies to Hitchcock's version of late preemption.[38])

7 Overlapping

Example 8

Two radioactive elements, atom 1 and atom 2, can decay, with atom 1 breaking down into atom X and particle Y and atom 2 breaking down into atom Z and particle Y.[39] Consider:

A_1 – atom 1 was brought to the site,

A_2 – atom 2 was brought to the site.

In fact, there was a decay right after that, and consequently:

C – a Y particle was present at the site.

In addition, atom X was present at the site, but no atom Z was present there. Consequently, it is clear that atom 1 decayed but atom 2 did not decay. Therefore, A_1 was a cause of C,[40] and A_2 was not a cause of C. The challenge is whether an analysis that focuses on intermediate events can handle cases with seemingly no intermediate events, such as actual radioactive decay or the hypothetical radioactive decay case here.

Moving to our analysis, A_1 and A_2 each bear ab initio probability increase to C, with no reversers. But consider F:

F – atom 2 didn't decay.[41]

Given F, it makes no probabilistic difference to C whether atom 2 was brought to the site or not. Hence F is a screener for A_2 and C, and, further, a stable screener. There is no ab initio probability increase of A_2 and F and no non-null increasers; hence A_2 is not a cause of F. Thus, F qualifies as a neutralizer for A_2 and C so long as it satisfies the constraint imposed on neutralizers in Section 5. The constraint specified there, Condition (8), is applicable to a neutralizer-candidate whose occurrence temporally reaches the upper edge of the occurrence time of the event specified in C. In general, we must make sure that this constraint is satisfied when the processes at hand are discrete. The restriction in question is the requirement that a neutralizer-candidate such as F here, for A_2 and C, not yield probability 1 for C; more precisely, F does not qualify as a neutralizer if:

$$\mathrm{P}(C/F.W_{A_2})=1$$

However, in our case, this restriction is clearly satisfied.[42]

Hence F qualifies as a neutralizer for A_2 and C. So A_2 is causally irrelevant to C and thus is not a cause of it, and this is indeed the right outcome. Yet there is no neutralizer for A_1 and C. In particular, an analogue of F (such as F_1: atom 1 did not decay) is not a neutralizer for A_1 and C since F_1 does not hold in our case. So A_1 is a cause of C, and this is, again, the right outcome.

8 Discrete electron levels

Example 9[43]

An electron was elevated to level 3 (A) (Figure 10.1). From there, it could descend directly to level 1, or first to level 2 (both with substantial probabilities) and then to level 1. (It could also stay at level 3 or descend directly to level 0, both with very low probabilities.) In fact, it descended to level 2 (B) and then it descended to level 1 (within a certain time interval) (C). But the descent from level 2 to level 1 was very unlikely. So the fact that it descended from level 3 to level 2 lowered the probability of its descent to level 1. Yet the fact that it descended from level 3 to level 2 was a cause of the fact that it descended to level 1. So B lowered the probability of C, even though intuitively B was a cause of C.

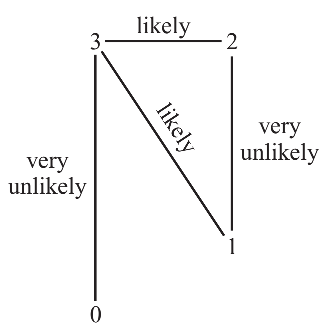

Figure 10.1

There is indeed ab initio probability decrease of B to C. For B to come out as a cause of C, on the above analysis, there must be an increaser. This raises a challenge for probabilistic analyses of cause, in particular for the above analysis, since direct transitions of electrons between energy levels presumably don't make room for intermediate events. However, in the context of approaches that concentrate on analysing physical causation as exemplified in physical phenomena (such as Salmon's or Dowe's), the physical features of the phenomena employed should be brought to bear. When an electron descends from one energy level to another, its energy is lowered, and conservation of energy requires that this lost energy doesn't just vanish. In fact, such an electron emits a photon of a characteristic frequency. For our electron to have descended from level 3 directly to level 1, it had to emit a photon of frequency X. However, it emitted a photon of frequency Y while descending from level 3 to level 2, and another photon of frequency Z while descending from level 2 to level 1. So consider:

E – no photon of frequency X was emitted.

Event E implies that the electron did not descend directly from level 3 to level 1. Thus, given E, the fact that it descended from level 3 to level 2 increased the chance that it reached level 1, and E is thus an increaser for B and C:

(9) $P(C/B.E.W_B) > P(C/{\sim}B.E.W_B)$

There is no reverser, hence E is a stable increaser, and there is no neutralizer for B and C. Hence B is indeed a cause of C, which is the right outcome.[44]

Wesley Salmon suggested difficult cases for probabilistic causation of a related sort, which he claims to be cases without intermediaries.[45] Thus, he writes: '… there are no further facts that are relevant to the events in question' (1984: 201). This sort of example played a major role in Salmon's conclusion: 'Thus, it seems to me, we must give serious consideration to the idea that a probabilistic cause need not bear the relation of positive statistical relevance to the effect' (1984: 202).

However, one might insist that we should still attempt to apply the account not to a physically viable case of trajectories of electrons, but to the above fictitious case construed as involving no intermediate events such as photon emissions during a descent of the electron. Yet, in this case the analysis still applies, since consider:

H – the electron didn't descend directly to level 1.

Given H, B yields probability increase (since the probability of C given ${\sim}B.H.W_B$ is 0 whereas the probability of C given $B.H.W_B$ is just low), and thus H is an increaser for B and C, and indeed a stable increaser, and B comes out as a cause of C. (Similarly, one could also employ the increaser: K – the electron didn't remain at level 2.)

Salmon was thus not quite right when he stated that ' … it appears that there is

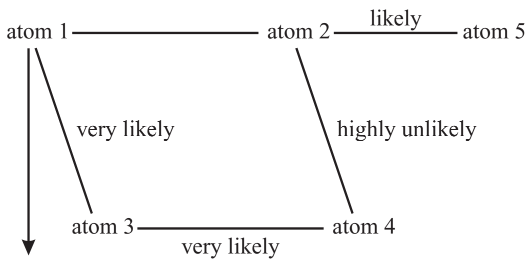

Figure 10.2

no way, even in principle, of filling in intermediate events' (1984: 201), so long as one is sufficiently liberal regarding what counts as being intermediate, and not insist that *A* be strictly before *E* and *E* be strictly before *C* for *E* to count as an intermediate event for *A* and *C*, and also, last but not least, so long as one selects intermediate events in the right way, as I have attempted to do in Sections 1–4.

Finally, consider Example 10, about radioactive decay, which is also directed against probabilistic accounts of cause.

Example 10

In it atom 1 can decay directly into either atom 2 or atom 3 (or it can also decay in another way). In fact:

A – atom 1 decayed into atom 2.

Atom 2 decayed into atom 4, which was highly unlikely. Thus:

C – atom 4 was present.

A was a cause of *C* (see Figure 10.2 below). But *A* decreased the probability of *C*, since ~*A* yields, with very high probability, that atom 1 decays into atom 3, and then further decays with a high probability into atom 4.

Again, on the above analysis, consider:

E – the descendant of atom 1 didn't remain in the form of atom 2 nor did it decay further into atom 5 (nor did it disintegrate).

E is compatible with *A* as well as with ~*A*. Given *A*, in the above example, *E* yields probability 1 of a decay into atom 4. Given ~*A*, atom 1 either decayed into atom 3 and stayed in that form or, from it, with high likelihood, descend into atom 4, or else decayed but not into atom 4 via the other path designated in the diagram by the vertical offshoot from atom 1. So *E* is a stable increaser for *A* and *C*, and there is no neutralizer.

Consider now a variation of the example, in which we rule out the third possibility (the vertical one in the diagram) of decay out of atom 1. In this case, there is still probability less than 1 of C given $E.{\sim}A.W_A$, yet there is probability 1 of decay into atom 4 given $A.E.W_A$. Thus, is in this version too, E is a stable increaser. (One can, in addition, vary the example still further by ruling out the possibility of a decay of atom 2 into atom 5. Then the following F will suffice as a stable increaser: F – the descendant of atom 1 didn't remain in the form of atom 2 (nor did it disintegrate), and so will the following G: G – there was no decay of atom 3 into atom 4.)

Notes

1 For the full analyses, see Kvart 1997, Kvart 2001a and Kvart 2003 (available also in my home page: http://socrates. huji.ac.il/Prof_Igal_Kvart.htm).

2 Jonathan Schaffer raised challenges of the first sort; see the discussion in section 7. Phil Dowe raised challenges of the second sort; see the discussion in section 8.

3 In this paper I deal only with token causes and token causal relevance. I employ narrow event individuation, Kim-style (which is called for since token events are to have probabilities), appropriately extended. Note that the notion of an event, as used here, is quite liberal – including, for example, states, processes, omissions, and so on – and is akin to Mellor's facts. Paradigmatically, a narrowly individuated event consists of an object, a property and a pertinence time. We allow ourselves, here and elsewhere, to use A, B, C and so on ambiguously as names of sentences as well as names of the events specified by these sentences (narrowly individuated). For further details about the kind of narrow individuation employed here, see Kvart (2003).
I also assume here a non-relativistic framework. However, since in general the order of temporal priority is assumed here, this assumption can be construed as having the effect that the notion of cause is frame-dependent.

In a Markovian world, using the world history up to A is tantamount to using a world state during a tiny interval just before A. I prefer the former since, in checking for intuitions in specific examples, an epistemic correlate must be employed, and the epistemic correlate of the world-history up to A is obviously preferable.

4 The problem arose both in cases of generic cause and in cases of token cause. Recourse to cause transitivity partly addresses this issue, but only regarding the necessary condition for being a cause, where transitivity might be taken to play a role. However, as I have argued repeatedly and as have other writers, cause transitivity is not valid. For the earlier general argument of mine against cause transitivity, see Kvart 1991b, where I argued that causal relevance is not transitive due to diagonalization. But causal relevance is a necessary condition for being a cause, and this in turn gives rise to the non-transitivity of cause. And indeed, the notion of cause is vulnerable to diagonalization as the notion of causal relevance is, and thus cause is also non-transitive. For a counterexample of mine against cause transitivity pertaining to a distinct and different source o*f* failure, see Kvart 1997: section 5; and Kvart 2001b: section 1. Michael McDermott (1995) offered the dog-bite counterexample, which I argued, however, is an instance of diagonalization (see Kvart 2001b: section 1; forthcoming c: section 12). Other successful counterexamples against cause transitivity have been offered by Ned Hall and Hartry Field (see Hall 2000, forthcoming; Lewis 2000).

The problem of causes that lower the probability of their effects, which has sometimes been handled by recourse to transitivity, is here handled by the requirement of a stable increaser. The corresponding problem of non-causes despite probability increase is handled, on my account, in one kind of case, by the requirement of a stable increaser (and not just ab initio probability increase), and in another kind of case, by the requirement of causal relevance, which involves another kind of intermediate events – causal-relevance neutralizer. This problem, in the latter case, was observed also by Peter Menzies (1989), who noted the failure of raising the probability of the effect as a sufficient condition for being a cause in view of causal chains that are cut off; see also Kvart 1991a, where I stressed the importance of causal relevance for being a cause (and for counterfactuals).

5 For detailed arguments, see Kvart 1991a, which shows that causal relevance (and thus cause) strongly depend on the intermediate course (cause of course requires causal relevance). See also Kvart 1997: section 8.

6 Assume that *x*'s financial state at t_A can be summarized by *Z*. Strictly speaking, *C* should be read as: *x*'s financial position was significantly better than *Z*.

7 Assume that *x* answered the questions in their alphabetical order.

8 Menzies noted such insufficiency only in cases when causal routes are cut off (see note 4). For a detailed analysis of this notion of cause, understood in terms of ex post facto probability increase and employing the notion of an increaser, see my 1997 'Cause and Some Positive Causal Impact', sections 6–12.

9 In discussing intermediate events, I consider only actual events. Intermediate events, as I noted, must belong in the (t_A, t_C) interval. Pertinence times and occurrence times are, howver, distinct. t_K, the time interval to which any given true, factual sentence *K* pertains, may extend beyond the occurrence time of the event specified in it, the *K*-event. The occurrence time of intermediate events, whether reversers or neutralizers, must of course not exceed the upper edge of the occurrence time of the *C*-event or the lower edge of the occurrence time of the *A*-event.

10 I have used the notion of a strict increaser in my earlier writings, but the term 'stable increaser' seems better mnemonically. I will not discuss here the possibility of further constraints on which events should count as increasers or decreasers.

11 This will be at the very least a convenient position to adopt throughout for the purpose of exposition and motivation. We need not take a position at this juncture regarding Humean supervenience of chances on facts.

12 Given causal relevance. Thus, if you will, consider in this case an empty intermediate event a null increaser.

For a detailed analysis of this notion of cause, understood in terms of ex post facto probability increase and employing the notion of an increaser, and for more on the notion of an increaser, see Kvart 1997. I offered this account of stable increasers as a solution to the problem of causes that lower the probability of their effects, a well-known problem in probabilistic accounts of causation, as well as to the problem of events that raise the probability of later events that are nonetheless not their effects, in Kvart 1994. When I wrote the latter paper I was not yet ready to propose and stand behind the thesis that the presence of a stable (strict) increaser (on top of causal relevance), or of some positive causal impact (which, in my view, are equivalent), amount to the relation of being a cause. (I thus proceeded to use, instead, the stable increasers analysis as a stepping stone to the more complex notion of overall positive causal impact.)

13 I assume that the avalanche occurred as soon as the dog was released. If there is one causal-relevance neutralizer, there may be more. For the purpose of characterizing causal relevance, all we need in characterizing a neutralizer is to specify when an event secures the absence of any would-be causal chains from *A* to *C*. It matters little for the purpose of the above analysis of causal relevance whether such a specification captures as a neutralizer the event that does the cutting so long as it secures the presence of one neutralizer or other.

14 I assume here that the avalanche didn't affect the chance that the second enemy hit *x*.

15 This notion should be understood in terms of an ab initio outlook as of t_A.

16 The issue of whether or not there is a neutralizer arises only if there is probabilistic relevance, which is assumed here. (6) presupposes here, for simplicity, a null differentiator, rendering E a screener simpliciter. If there is a non-null differentiator, there may be a screener *for it*. The requirement below that a neutralizer E be a stable screener yields that E screens off A from C simpliciter or in the presence of any non-null differentiator.

I have previously used also the term 'blocker', rather than 'screener'. I use these two terms interchangeably.

17 More precisely (see note 16): If A is probabilistically relevant to C, a stable screener (for A and C) is an intermediate E, fulfilling (6), such that no intermediate F fulfils (7).

18 Assume further, that the chance of any other change in either pipe, given $E.F.A$ vs. given $E.F.{\sim}A$, is the same.

19 At least until a sufficiently short time before C. (I implicitly assume, for E's being a causal-relevance neutralizer, that E was independent of A.)

20 More generally, true informational expansions of a stable screener are also stable screeners. A conjunctive expansion of a stable screener E is an event $E.F$, for any intermediate F. Various true informational expansions of a causal-relevance neutralizer also do the job of a causal-relevance neutralizer. If there are causal-relevance neutralizers, there are lean ones. So we may focus our attention on lean stable screeners in our attempt to characterize causal-relevance neutralizers suitable for securing causal irrelevance. In securing causal irrelevance, we need only to secure the presence of just one event fulfilling the role of a causal-relevance neutralizer.

If a certain actual event E in fact secures the neutralization of the causal relevance of A to C, but this is not revealed ab initio – that is, from the usual perspective of the upper end of W_A, or in other words, given just W_A – then some true informational expansion of E secures it ab initio as well by securing that various would-be causal threads are neutralized. Therefore, if causal irrelevance is attested to by an actual event E that in fact secures the neutralization of the would-be causal relevance of A to C, but E is not a stable screener, such causal irrelevance will also be attested to by a causal-relevance neutralizer that is a stable screener (in particular, by some true conjunctive expansion of E). Hence, in specifying the conditions for causal-relevance neutralizers, we can confine our attention to stable screeners, without loss of generality.

Consequently, we need not be overly precise here about the notion of a lean screener. All we need is to characterize an event that secures the neutralization of would-be causal relevance. If there is one, there are many, and for establishing causal irrelevance it does not matter which one we pick. Hence we may characterize causal-relevance neutralizers in a way that brings out the fact that they serve that role and yet do not include ingredients that are immaterial to that role.

21 Since the world is Markovian, any intermediate full world state is a stable screener for any event that precedes it vis-à-vis any event that occurs after it. We can consider such states *global* screeners.

22 Assume that all the parameters of the bullet at t_1 that are pertinent to its future course are specified in E_1.

23 Accordingly, there are infinitely many other such exhaustive channellers along this trajectory. Such exhaustive channellers can also be specified in cases where there is more than one causal route from A to C, for example, as conjunctions of such *local* stable screeners vis-à-vis a given route. For further details, see Kvart 2001a: esp. section 5.

24 Note 13 rules out that the dog's charging forward triggered the avalanche. The very release of the dog, we assume, yielded no vibrations of significance. In any case, there are alternative neutralizers, other than the above E, such as E': the dog didn't come close to the target (until t_C). Implicit in Example 3 on p. 139 about the pipe and the tap was that A was causally irrelevant to E there. In earlier writings I used the term '*crn*' as an abbreviation for 'causal relevance neutralizer'.

25 And if there is one, there are many. One central feature of this analysis of cause is that in preemption cases (early or late) the preempted cause is not a cause since it is causally irrelevant to the effect, a feature that is secured by the presence of a causal-relevance neutralizer. For a detailed analysis of how this approach handles such cases, see Kvart (forthcoming d: sections 8 and 9, and this volume: section 6).

Note, however, a caveat to this analysis. Causal relevance and the presence of a stable increaser are necessary, but not quite sufficient, for being a cause. It must be further ascertained, when checking for causes, that there is no purely negative causal relevance despite the presence of a stable increaser, which can be present if the route of the would-be positive causal relevance (indicated by a stable increaser) is neutralized by a *positive relevance neutralizer*. This case can be analysed with the notions used here for analysing causal relevance; see Kvart (forthcoming b). I ignore this complication below.

26 See Kvart 2001a: section 8, or, in a brief version, Kvart 2003: section 5. Note, however, that the same pattern holds not just under the assumption that time is continuous (which seems essential for physics as we know it), but also under the assumption that time is merely dense, both of which allow for infinite regress. If time is discrete, then still causal relevance amounts to the absence of a neutralizer (given probabilistic relevance); but since all such sequences are finite, causal relevance is no longer exhibited by the presence of infinite regress. In such a case one must be especially attentive to the constraint on neutralizers that pertain to the upper end of the occurrence time of C, and in particular in a case of E_n that occurs one temporal unit after t_A. Such a constraint is introduced on p. 172 in section 5, and in section 7 we consider in detail its application to cases of this sort regarding radioactive decay (albeit fictional ones).

27 For a more detailed presentation of the above analysis of cause, see Kvart 1997: sections 6–12. For a more detailed presentation of the above analysis of causal relevance, see Kvart 2001a: 59–90. For a related treatment using the above account for an analysis of the thirsty traveller puzzle, see Kvart 2002. All of these papers are available also in my home page: http://socrates.huji.ac.il/Prof_Igal_Kvart.htm.

28 For more on purely negative causal relevance, see Kvart 2001b: section 4, 1986: section VIII. Recall also that we allow ourselves, here and elsewhere, to use A, B, C and so on ambiguously as names of sentences as well as names of the events specified by these sentences (narrowly individuated).

29 Kvart 1997: section 6.

30 This example is taken from Kvart 2001b: section 4.

31 See Kvart forthcoming a: section 8.

32 Under narrow event individuation, the framework within which we operate here, an event is paradigmatically specified by an object, a property (or a predicate) and a specified occurrence time – that is, a temporal interval qualified by a temporal quantifier indicationg the fragility of the event in question. Lewis's notion of fragility focused on modal aspects, which are important for a counterfactual account of cause. But for the probabilistic approach to cause advanced here, the modal aspect is not as important.

33 Kvart 2002: section 8.

34 I take this to be the brunt of a challenge posed by Murali Ramachandran.

35 Another neutralizer for B and C is: G – Billy's rock did not hit the window until the time it was shattered (if it was shattered). G screens B off from C, and stably, and B lowers the chance of G (see also note 37), with no stable increaser, and thus B is not a cause of G. So G is a neutralizer for B and C, and indeed constraint (8) presented in section 5 is satisfied: the probability of C given G and W_B is less than 1. Hence B is causally irrelevant to C and thus is not a cause of it, and this of course, intuitively, is the right outcome.

The antecedent of G, in other words 'if it was shattered', implicitly pertains all the way up to T_3– (that is, the upper end of T_3). So the specified occurrence time of G, t_G, pertains up to T_3–. This feature of G is built into the specification of G; G must qualify as an intermediate event that screens off *B* from *C*. Thus, when we consider the equality that governs screening, that is: (6') $P(C/B.G.W_B) = P(C/{\sim}B.G.W_B)$ it is left open, given W_B, whether or not *C* is actual. If t_G– (that is, the upper end of the specified occurrence time of the *G*-event) were allowed to be earlier (at least sufficiently so) than t_C–, even if in fact it is later than the upper end of the actual occurrence time of the *C*-event, it would be left open, regarding equality (6') above, whether Billy's rock hit the window after t_G– and prior to the actual occurrence of the *C* event (again, insofar as equality (6') is concerned).

Yet one may argue that the upper end of the actual occurrence time of *G* may do even if taken to be earlier by an ε than that of *C*, and the same applies to *H*. If so, the edge condition is otiose insofar as this example, and presumably late peemption in general, are concerned.

36 Still another suitable candidate for a neutralizer for *B* and *C* is: *D* – Suzy's rock shattered the window before Billy's rock reached it (if the window was shattered). *D* doesn't yield *C* with probability 1, so it passes the edge condition (8). *D* screens off *B* from *C*, and *B* indeed decreases the probability of *D* with no stable increaser, and thus is not a cause of *D*. Note that selecting fragile *A* and *B* affects neither the analysis nor its outcome.

37 In Kvart 2003: section 8, I offered a somewhat different neutralizer for a late preemption case similar to the one discussed here but in a version in which the effect-candidate was temporally fragile (in it, 't_3' replaces '*T*3' in *C* here). It centred on a neutralizer that didn't pertain all the way up to the upper end of the occurrence-time t_3 of the *C*-event, since in that paper I did not prepare the ground for the edge condition for neutralizers. The solution offered here applies both to the fragile as well as to the non-fragile versions, and in section 5 the ground is prepared for neutralizers the upper ends of the occurrence times of which reach all the way to the upper end of the occurrence time of the effect. Thus, it provides for a general resolution of the general late preemption case.

However, depending on the circumstances and the chance distribution given W_B, *B*, especially in a sufficiently non-fragile version, might bear ab initio probability increase to *G*. In order to make the point in a simple way, consider, instead of *B*, a temporally fragile *B'*: Billy threw the rock at t_2 (where t_2 is instant-like, rather than a significant interval). Yet the circumstances might be such that throwing a bit later would yield a very high probability of throwing in the direction of the window with a sufficiently greater force, and thus a very high probability of arriving earlier and, consequently, a higher probability of hitting the window. In such a case, if not throwing at t_2 yields a high enough chance of throwing a bit later, *B'* may increase the probability of the above *G*. Thus the above *G* need not be a neutralizer for *B'* and *C*, although *G* screens off *B'*.

B', under such a scenario, seems to be an ab initio (as distinct from ex post facto) delayer of the fragile version of *C* here (that is, with t_3, instead of T_3), and it is probabilistically relevant to *C* in this case. But since it yields ab initio probability decrease to *C*, with no stable increaser, it is not a cause of *C*. (Yet it is a cause of the fact that Suzy's rock shattered the window. For reasons further supporting the last claim, see Kvart forthcoming c.

38 See Hitchcock forthcoming. In Kvart 2003: section 8, I treated Hitchcock's version without assuming that a neutralizer may temporally extend all the way up to the temporal upper edge of the *C*-event.

What is the occurrence time of a disjunction? The pertinence time of a disjunction is the smallest interval that includes the pertinence times of both disjuncts. But if one disjunct is false, it has no actual occurrence time, and then the actual occurrence time of the disjunction is that of the true disjunct.

Note that a variant of condition (8) will do as well: (8') The upper edge of the actual occurrence time of a neutralizer *E* for *A* and *C* is earlier than that of *C*, *if C* is actual. (8') is stronger than (8), and so if it obtains, it protects against the concerns that motivated the imposition of condition (8) in cases in which the upper edges of the specified occurrence times of E and C coincide and so its usefulness is apparent in the latter cases. (8') can be shown to hold regarding G of note 35, and so, given that, no extra attention is needed for ascertaining that (8) holds.

39 Jonathan Schaffer proposed an example of this sort in order to argue that probability-raising theories of causation, mine in particular, fail in view of processes with no intermediate events; see Schaffer 2000b. I chose the terminology of atom 1 breaking into atom X while emitting particle Y, and correspondingly for atom 2, rather than the terminology of atom 1 and atom 2 emitting particles X and Y and particles Y and Z respectively, in order to attempt to keep the example within the physically viable realm of radioactive decays where, it seems to me, it would have a greater force. In any case such a modification does not seem to diminish the force of the original example.

40 Note that had atom 1 not been brought to the site, a Y particle would not have been present at the site. We can also assume that $t_{A_1} = t_{A_2}$, and that atom X and particle Y were present at the site right after atom 1 reached the site.

41 Strictly speaking, a full specification of C must include a temporal specification, which in our case would most plausibly be instant-like. The time to which F pertains is the same as that of C (and same for F' of note 42).

Schaffer argued (private communication) that F is not an event but rather a process. But the notion of events (narrowly individuated) as used in the present analysis for causal relata is construed liberally, and covers processes, states, omissions and so on, and is more akin to the notion of a fact (see note 3). Thus, a concern about whether such an F is intuitively a process doesn't affect its being an event as this notion is employed in the analysis above.

42 Note that F yields a lower probability of C in comparison with $\sim F$, or equivalently: $P(C/F.W_{A_2}) < P(C/W_{A_2})$, and in our case: $P(C/F.W_{A_2}) < 1$.

Alternatively, instead of F, one could use as a neutralizer the following F': atom 2 was still present.

Although the restriction that a neutralizer E for A and C must yield that $P(C/E.W_A) < 1$ did not affect our treatment of this example in its original form, consider the following G: atom X was present. G would not do as a neutralizer for A_2 and C: although it is a stable screener for A_2 and C, it violates the above restriction. (It is important for Schaffer, regarding atom X and particle Y, that atom 1 cannot yield one without the other.) If there are two Y particles, then both atoms decayed, and both A_1 and A_2 are causes of C.

43 Dowe offered (private communication) an example of this sort, as well as of the sort of the next example about radioactive decay; see also Dowe 2000b.

44 $t_E = (t_B, t_C)$. Note that it is physically possible for the electron to remain, within a time interval that encompasses the time of C, at level 3, or to descend directly from level 3 to level 0, which is lower than level 1, so $\sim B.E.W_B$ is physically possible. $P(C/\sim B.E.W_B)$ is 0, whereas $P(C/B.E.W_B)$ is just very low.

Can the photon emissions serve as neutralizers? E, as is evident from (9), is not a screener for B and C. Consider F: a photon of frequency Y was emitted. But F yields B with probability 1, and thus F is not a neutralizer for B and C (recall the restriction at the end of Section 5). Similarly, consider G: a photon of frequency Z was emitted. But B raised the chance of G, and so B is a cause of G, and thus G is not a neutralizer for B and C either. (Consequently, $E.F$ yields B with probability 1, and B is also a cause of $E.G$.) Recall that the qualification mentioned in Section 5 on p. 173 was imposed for neutralizers. But this does not mean that reversers are altogether free of restrictions; see Kvart forthcoming b.

Note further that the assumption that the electron can descend directly to level 0 is not crucial to the applicability of the analysis. Suppose it is dropped; the above analysis still holds, with H being an increaser.

In addition to H, we can also use K: K – the electron did not remain at level 2.

$P(C/B.K.W_B) = 1$; $P(C/{\sim}B.K.W_B) < 1$, and so *K* is a stable increaser as well (the electron has certain probabilities to descend to lower levels, but it also has a non-0 probability to remain at a certain level over a certain period of time). Postulating $P(C/{\sim}B.W_B) = 1$ implies that the electron has probability 0 to remain at level 3 (over a certain period), whereas in a physically viable case, such a probability is non-0. The analogue, in a radioactive decay case, is the atom's having a non-0 probability of not decaying, determined by its half-life time. A constraint to the effect that $P(C/{\sim}B.W_B) = 1$ takes us beyond the realm of the physically viable, and may be construed as counter-nomological.

However, the above probabilistic analysis, as presented, is not unlimited in view of cases with probabilities 1, and in particular deterministic cases. Yet discrete-time cases, which do not allow for 'natural' intermediate events, do not seem to present *as such* a major impediment. One may deal with cases with probability 1 by using non-standard models, or by moving to qualitative probability. I pursue in particular the latter strategy for extending the above analysis to the deterministic case.

45 See Salmon 1984: 200–1. Salmon's example ignores such photon emissions in the descent of electrons to lower energy levels. Salmon is unlikely to have been impressed by the difference between facts and events regarding *E* (see the quotation in the text from his p. 201). Although Salmon's example is admittedly fictitious, Salmon cannot be reasonably construed as deliberately ignoring conservation laws.

I thank Itamar Pitowsky and Matar Wax for comments on portions of this paper.

Chapter 11

Prospects for a counterfactual theory of causation

Paul Noordhof

My aim is to defend a counterfactual analysis of causation against purportedly decisive difficulties raised recently, many rehearsed and developed further in this volume. Although some of the moves I will make are available to any counterfactual theory, my principal aim is to explain how a theory I outlined elsewhere can, with some adjustment and simplification for the purposes of discussion, deal with a range of problems (see Noordhof 1999 for original presentation of the theory). Specifically, I will be concerned with the issue of whether the semantics of counterfactuals can be characterized independently of causation (raised by Dorothy Edgington, this volume), the proper way to deal with the nontransitivity of causation (raised by Michael McDermott 1995 and Murali Ramachandran, this volume), and a collection of counterexamples to the idea that causation involves, at its heart, chance-raising (discussed in this volume by Helen Beebee; Phil Dowe; Doug Ehring; Chris Hitchcock and Michael Tooley, and by Jonathan Schaffer (2000a, 2000b)). Obviously, in defending my own counterfactual theory, I am also implicitly arguing that counterfactual approaches to causation in general have the resources to capture its important features. The ambiguity in the title thus accurately reflects the content of the present paper.

1 My theory (or at least part of it)

My central idea is that causes are events which, if they occurred independently of their competitors, would make the chance of their effects very much greater than, at the limit, their general background chance at the time at which the effects occurred via an actually complete causal chain. Hence I do not take causal claims to be essentially contrastive: true or false relative to whichever alternative scenario is had in mind (see for example Hitchcock (forthcoming)). I appealed to counterfactuals and various other notions in order to capture this idea. A simplification of my theory runs as follows:

For any actual, distinct events e_1 and e_2, e_1 causes e_2 if and only if there is a (possibly empty) set of possible events Σ such that

(I) e_2 is probabilistically Σ-dependent on e_1, and,

(II) every event upon which e_2 probabilistically Σ-depends is an actual event,
(III) e_2 occurs at one of the times for which $p(e_2 \text{ at } t) \geq x \gg y$.[1]

I define probabilistic Σ-dependence in the following way:

e_2 probabilistically Σ-depends upon e_1 if and only if:

(1) If e_1 were to occur without any of the events in Σ, then for some time t, it would be the case that, just before t, $p(e_2 \text{ at } t)$ generally around x,
(2) If neither e_1 nor any of the events in Σ were to occur, then for *any* time t, it would be the case that, just before t, $p(e_2 \text{ at } t)$ generally around y,
(3) $x \gg y$.

Let me offer a few preliminary comments in the way of explanation. Other features of the proposal will become clearer when I turn to some of the problems that have been raised for approaches like mine.

Talk of Σ-dependence is a mechanism by which to take away competitor possible causes, for example, in cases of preemption or over-determination. When there is a preempted or back-up chain, it need not be true that what we might intuitively count as a cause would be necessary in the circumstances, or significantly raise the chance of an effect. However, it would still be true *if* events in the preempted chain did not occur (that is they were put in the Σ-set). The definition of probabilistic Σ-dependence defines a notion of chance-raising conditional upon the events in the Σ-set being absent. The time of assessing the chance of the putative effect is just before the occurrence of the effect. If we put an event of the preempting causal chain into the Σ-set, then the preempted chain can run to completion. That often means that there will be an event, which didn't actually occur, occurring in these changed circumstances. We don't want to conclude that the preempted chain contains causes of the putative effect. After all, it was preempted. Clause (II) rules this out by insisting that there should be no non-actual events upon which the putative effect probabilistically Σ-depends. We just saw that in the case of the preempted chain this may well not be the case. Of course, the preempted chain may not be filled in. But in that case, it will not be true that the chance of the putative effect assessed just before its occurrence is raised by the occurrence of the putative, in fact preempted, cause. '$x \gg y$' should be read as x is proportionately very much greater than y. This does not mean that the probability of x should be high. Clause (III) and assessing the chance of e_2 occurring at a time t ensures that we are not just considering whether the chance of the putative effect, e_2, is raised but more importantly whether it is raised by e_1 at the time that e_2 actually occurred (for details and further discussion see Noordhof 1999: 108–14).

I appeal to the idea of chance-raising because I think it is plausible that there are indeterministic causes. Some have denied that appeal to chance is necessary (such as Ramachandran (1997), Barker (this volume: 120–4, 132–4)). Simple counterfactual dependence is all that is required. I remain unconvinced (see Noordhof 1998a: 459–60).

It seems to me that those who deny that appeal to chance is necessary face one challenge and have one unargued commitment. The challenge is to explain an asymmetry. It is alleged that, in every putative case of indeterminism, even though there is not a sufficient cause of a certain effect, there will be conditions which, together with the absence of the putative cause, will be sufficient for the effect to have no chance of occurring.[2] Why should indeterminism always take this structure? The asymmetry can seem plausible if care is not taken over the specification of the conditions for the absence of the effect. I don't deny that one can identify conditions sufficient for the effect to have no chance of occurring. Consider a radioactive isotope of a chemical element such as radium or uranium. The isotope will have a chance of decaying whether it is bombarded by subatomic particles or not. Can we identify conditions under which there will be no chance of the isotope decaying? It is clear that we can. There will, of course, be no chance of decay if there is no radium or uranium isotope present. But notice that, in this case, we are citing something which partly constitutes the effect. Mention of the absence of bombardment is redundant. The chance of decay would be zero whether or not there is bombardment. The issue is whether we can always identify something which requires the absence of the putative cause *as well*. Barker's discussion of this issue in terms of the example mentioned cites the absence of any energy state in the isotopes (Barker, this volume: Sections 1, 4). I think it is reasonable to wonder whether, in that case, mention of the absence of the bombardment is essential to the chance of decay being reduced to zero. It seems clear that it is not. In which case, we would not have the required counterfactual dependence of decay upon bombardment. We can fight over particular cases but, to have an entirely general theory, appeal to chance-raising and -lowering seems inevitable.

A key idea of my proposal is that causes are not just chance-raisers of an effect at *a* time relative to Σ. Rather they make the effect event very much more likely to occur than it would at that time *or any other time* had the cause been absent. The reason for taking this line stems from consideration of hasteners and delayers. Jonathon Bennett's nice example of the April rains and the forest fire illustrates the point (Bennett 1987: 373–4). The forest would have been dry in May and caught fire if the rains had not come in April. However, by June, the forest had dried out again and there was a forest fire. If the fire had occurred in May, there would have been no fire in June because there would have been nothing to catch fire.

I assume that it is possible to delay a particular token event from occurring, in the present case, the forest fire. It is not plausible to claim that, in every putative case of delay, we really have the causation of another event of the same type a little later. Again, we can squabble about particular cases but the general point stands. The difficulty is that such cases of delaying do not seem to be causings of the delayed event even though they are causings of the event occurring at a particular time. The April rains did not cause the forest fire even though they were among the causes of the fire occurring in June. If I had characterized probabilistic Σ-dependence simply in terms of chance-raising at a time, the April rains would have come out as a cause. The reason they do not is that April rains do not make the forest fire very much more likely in June than it would have been at any other time, for instance, in May.[3]

2 Counterfactuals and circularity

The first type of objection I'm going to consider is supposed to be entirely general. If it is correct, it works against all counterfactual theories, both those theories which appeal simply to counterfactual dependence, or its ancestral, to capture causal dependence and to theories such as mine which appeal to counterfactuals to characterize the idea of chance-raising. The basic charge, outlined most impressively by Dorothy Edgington in the present volume, is that the proper semantics of counterfactuals must mention causal facts. Hence, counterfactuals cannot be used to provide an analysis of causation. I shall expound the problem within David Lewis's framework though, as Edgington notes, it is not specific to it.

Lewis suggested that the truth conditions of counterfactuals should be given as follows:

> A counterfactual 'If it were that A, then it would be that C' is non-vacuously true if and only if some (accessible) world where both A and C are true is more similar to our actual world, over-all, than is any world where A is true but C is false.
>
> (Lewis 1979: 41)

Lewis's criteria for assessing the similarity between possible worlds are as follows:

> (A) It is of the first importance to avoid big, widespread, diverse violations of law.
> (B) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
> (C) It is of the third importance to avoid even small, localized, simple violations of law.
> (D) It is of little or no importance to secure approximate similarity of particular fact, even in matters which concern us greatly.
>
> (Lewis 1979: 47–8)

Edgington argues that (D) must be adjusted so that we retain approximate similarity in causally independent facts after the antecedent.

One case Edgington uses to illustrate her point runs as follows. A general activates a missile which, as a result, has a 25% chance of firing. As it happens, it does fire and hits its target, the village of two people coming home from the fields. A little earlier, the villagers were delayed trying to help a sheep stuck in a ditch. When they arrive at the village to find it destroyed, one remarks to another:

> (1) If we hadn't noticed the sheep stuck in the ditch, we would have been killed by the missile hitting the village.

If we followed Lewis's similarity weighting, we ought to judge that (1) is false. If we consider the development of the world after the slight change required for the villagers to fail to notice the sheep, then it is not determined that the missile should fire. It only has a 25% chance of firing. So rolling on the world in line with the laws would not imply that the missile would fire and, indeed, would suggest that probably the opposite is the case. Hence (1) would be false. On the other hand, if we take up Edgington's idea that we should bring forward causally independent fact with the laws, then it is clearly the case that the villagers' noticing the sheep is independent of the firing of the missile. So we can keep the missile's firing fixed as part of the circumstances we consider in assessing (1). Now we get the right verdict: the truth of (1).

Edgington's argument raises a number of issues. The first point to make is that the proponent of the counterfactual theory of causation does not have to eschew the specification of (D) that Edgington proposes. We can use counterfactuals understood via Lewis's similarity weighting to capture preliminary facts of causal independence in terms of which we can then provide a new similarity weighting which adverts to these causally independent facts. We can then appeal to our new counterfactual judgements to characterize further causally dependent facts, and so on. For instance, in Edgington's case, if we start with Lewis's similarity weighting, we will arrive at a range of facts we can bring forward to consider whether the firing of the missile is causally independent of the villagers noticing the sheep stuck in a ditch. We can now consider the counterfactuals with this new assumed background to see whether there are any causally dependent facts that need to be weeded out, and so on. Our judgement that the firing of the missile is causally independent of the villagers' noticing the sheep stuck in the ditch is true if there is no new similarity weighting derived from application of the procedure just outlined that demonstrates that the firing of the missile is causally dependent upon the villagers' powers of observation.

This mechanism is also revealed in assessing the following counterfactual:

(2) If we hadn't noticed the sheep stuck in the ditch, we would be able to help you take the table upstairs.

Here our slightly lazy villagers are staying with a friend who needs some furniture moved to make up a bed for the night in the living room (since their village and homes are destroyed). They are claiming tiredness due to their earlier exertions. I take it that, although there is an interpretation under which (2) is true, a legitimate reply would have been: No, you wouldn't have been able to help because then you would have been dead. Our uncertainty over this counterfactual stems from a clash between the verdicts given by Lewis's similarity weighting and Edgington's proposed adjustment. By the preceding considerations, the death of the villagers is causally dependent on the truth of the antecedent. Hence their survival cannot be brought forward. This is so even though the preliminary judgement would have been that it is causally independent. Yet the destruction of the village remains causally independent of their noticing the sheep. So it should be brought forward. Lewis's similarity weighting makes (2) true, Edgington's adjustment makes it false.

Second, although I am sympathetic to Edgington's characterization of the relevant approximate similarity mentioned in (D), it is not clear that she has made the right diagnosis. She asserts that we should bring forward only facts causally independent of the antecedent. Instead, it seems to me that we should bring forward only *probabilistically independent* facts. A case she discusses, one originally put forward by Pavel Tichy, makes the point (Tichy 1976). Suppose that Fred takes his hat 90% of the time when it rains but only 50% when the weather is fine. In the present case, he takes his hat and it is raining. Now consider the following counterfactual:

(3) If it had not been raining, he would have taken his hat.

I think that intuitively this counterfactual is false. However, that is not because we suppose that its failing to rain would cause him not to take his hat. It may well be that the causal chain between it not raining and him not taking his hat fails to complete. Edgington's proposal would mean that we were *required* to include, as part of the context, him taking his hat. The reason why we should not take this as part of the context is that him taking his hat is not probabilistically independent of whether or not it had been raining. Inclusion or exclusion does not depend upon whether there is, in fact, a causal relationship. If that is right, there is no threat to attempts to characterize causation in terms of counterfactuals.

A natural way of capturing probabilistic dependence is to focus on whether the following held:

(PD1) If e_1 were to occur, it would be the case that, at t, $p(e_2)$ generally around x.
(PD2) If e_1 were not to occur, it would be the case that, at t, $p(e_2)$ generally around y.
(PD3) For any time of assessment t, $x = y$.

If not, then e_1 and e_2 are probabilistically dependent and hence e_2 should not be brought forward as part of the context in which we assess counterfactuals involving e_1's occurrence or failure to occur. As before, we would begin by taking these counterfactuals to be assessed by Lewis's similarity weighting and reiterate this procedure for the preliminary judgements concerning what is probabilistically independent. Obviously this proposal will require further assessment.

3 Transitivity and chance-raising

Causation is not transitive, contrary to some recent opinion. A classic example, due to Michael McDermott, is the dog-bite case (McDermott 1995: 531–2). Jo is about to trigger a bomb by pressing a button with his right finger but a dog bites it off. Hence Jo triggers the bomb by pressing a button with his left finger. The dog bite is a cause of Jo pressing a button with his left finger. Jo pressing a button with his left finger is a cause of the bomb exploding. Yet, the dog bite is not a cause of the bomb exploding.

Any counterfactual theory of causation which takes the ancestral of counterfactual dependence to be distinctive of causation has a potential problem with cases like this. So it seems that we shouldn't. However, there are considerations which also seem to point in the other direction, to my knowledge identified first by Murali Ramachandran (Ramachandran 2000: 310–11). Suppose that we consider an indeterministic causal chain made up of events e_1, e_2, e_3 and e_4 in succession. Each has a certain chance of occurring anyway even if the prior event in the causal chain did not occur. If that's right, then an obvious counterfactual formulation of chance-raising is in trouble if it is applied to cases of mediate causation. It is tempting to talk of the chance of a putative effect being 'at least' x if a cause is present and 'at most' y when a cause is absent to deal with the possibility of slight variation in circumstances across the closest worlds in which the antecedent is true. More specifically the temptation is to appeal to the counterfactuals:

(≥) If e_1 were to occur, it would be the case that $p(e_4) \geq x$,
(≤) If e_1 were not to occur, it would be the case that $p(e_4) \leq y$ (see for example Noordhof 1999: 104–5, and developments on pp. 108–15).

Unfortunately, if we suppose that the putative cause, e_1, is absent, then there is still a chance that one of the intermediaries will fire, so setting off the rest of the chain, and hence the chance of e_4 will be *at most* the chance it would have had if the cause had been present after all (assessed just before e_4). Therefore, e_1 will not raise the chance of e_4. Assessing the chance of the effect just after a putative cause, e_1, rather than just before the putative effect, e_4, does not resolve matters. e_4 might spontaneously occur at the time of e_1 and hence e_4's probability just after the time of e_1 will be 1. Once more we have no chance-raising.

It is at this point that taking the ancestral of counterfactual dependence to characterize causation becomes tempting. Suppose we stick with assessing the probability of putative effects just before their occurrence (if they occur and otherwise at the time of their actual occurrence). We can appeal to counterfactual dependence to characterize immediate causation since there are no indeterministic intermediaries to occur spontaneously. The ancestral of counterfactual dependence serves for mediate causation. Problem solved (even here, matters are not simple: see Ramachandran, this volume: 156–9). But now we are back with transitivity. That's the dilemma.

There seems to me to be three plausible ways to proceed. First, we might challenge the claim that if e_1 had not occurred, one of the two intermediaries e_2 and e_3, still might have occurred so giving rise to our problem. I canvassed this proposal in Noordhof 2000: 318–21). Although I am dissatisfied with the details of the proposal set out there, I remain convinced that the correct similarity weighting for counterfactuals would have this upshot. However, I do not have the space for a proper discussion of this issue and it involves some contentious matters. Second, we might appeal to the ancestral of counterfactual dependence and then place an additional condition to characterize when e_1 is a cause of e_4. For instance, Ramachandran suggests that this can be done by requiring that e_1 raise the chance of e_4 (assessed just after e_1) in the vast majority of (not necessarily actual) circumstances *given that neither e_4 nor any of the intermediaries occur until* ***after*** *(but not immediately after) the actual time of e_1's occurrence* (see Ramachandran this volume: 158). The difficulty with this suggestion is that one cannot assume that causal chains will run at the same rate in different circumstances. In which case, supposing e_1, e_2, e_3 and e_4 to be at the end of a longer sequence which is running more slowly, we might have a situation in which e_4 spontaneously occurs after (but not immediately after) the actual time e_1 occurred and yet, in the changed circumstances, this is just before e_1 occurs. If that happened, the chance of e_4 would be 1 and we still wouldn't have e_1's occurrence raising the chance of e_4's occurrence.

The third plausible way to proceed is to abandon, as I have done, the 'at least' and 'at most' formulation. In its place I have talked about values that p(e_2 at t) would *generally* have if e_1 is present or absent. In saying that p(e_2 at t) would generally be around x or y we don't have to take into account the exceptions mentioned above. The difficulty with this proposal is how we should understand 'generally' given that the domain of close possible worlds in which a certain antecedent holds may be infinite. We can't talk of the majority of the worlds being ones in which p(e_2 at t) has a certain value. There are no majorities in infinities. Instead, we have to allow that 'generally' adverts to a certain high probability that, if we were selecting from the relevant set of possible worlds, we would obtain a world in which p(e_2 at t) is around x (or y). This probability should not be understood in terms of proportions of possible worlds but would, instead, be based upon the character of the closest possible worlds in which the antecedent is true, specifically, the varying particular matters of facts and slightly varying probabilistic laws in these worlds. Obviously for the response to be fully satisfactory, more needs to be said about how this high probability is obtained. Nevertheless, it seems to me that there are no objections in principle to such an approach and no particular difficulties that don't already attend the rejection of frequentist approaches to probability.

With this indication of work for the future in place, let me explain how my own account avoids making causation transitive by briefly discussing the dog-bite case. The key conditionals are:

(4) If the dog bite were to occur without any of the events in Σ, then the chance of the explosion would be generally around x.
(5) If neither the dog bite nor any of the events in Σ were to occur, then the chance of the explosion would be generally around y.

It seems clear that, if nothing is put in Σ, x would not be very much greater than y and so the dog bite would fail to come out as a cause (which is what we want). The concern might be that there is something we might put in Σ where this is not the case. But this seems unlikely. For instance, consider an assignment to Σ of the right hand pressing. We would then get

(6) If the dogbite were to occur without the right hand pressing the button, then chance of the explosion would be generally around x.
(7) If neither the dog bite were to occur nor the right hand pressing the button then the chance of the explosion would be generally around y.

It seems obvious that x would still not be very much greater than y. On the assumption that Jo is trying to explode the bomb, we should assume that if he had not pressed the button with his right hand, he would have with his left. But, assuming indeterminism, the dog bite makes it less likely or, assuming determinism, at least not more likely for the explosion to occur. Hence x would not be very much greater than y.

4 Objections to taking causation as chance-raising

My account is committed to causal processes involving chance-raising in some specified circumstances (roughly, those in which the events in the Σ-set are removed). However, there is a range of cases that has been put forward in the present volume and elsewhere which seem to throw into doubt the idea that causation is linked to chance-raising. The purpose of the present section is to go through these cases and explain how my account yields the correct verdicts regarding whether or not causation has taken place.

The *Dowe–Salmon Radioactive Decay Case* involves an atom I which can decay by two possible routes to atom IV: via atom II and via atom III. The transition probabilities are as follows.

Atom I to atom II	0.5
Atom II to atom IV	0.1
Atom I to atom III	0.5
Atom III to atom IV	1

In his discussion of the case, Phil Dowe invites us to suppose that time is discrete and that there are no intermediaries. There is just decay from one atom to the other (Dowe 2000b: 33–40; this volume: 32–4). Suppose that, in fact, the decay to atom IV occurs via atom II. This is clearly the route which made decay to atom IV less probable. Nevertheless, Dowe argues, we suppose that the decay of atom I to atom II is a cause of its further decay to atom IV. So we have chance-lowering with causation (see also Tooley, this volume: 107–8).

My proposal would get this verdict by putting the possible event of atom III decaying to atom IV in Σ. In which case, we would be considering the following conditionals:

(8) If there were decay from atom I to atom II without decay from atom III to

atom IV, then for some time t, it would be the case that, just before t, p(decay to atom IV at t) is generally around x,

(9) If there were neither decay from atom I to atom II nor decay from atom III to atom IV then for *any* time t, it would be the case that, just before t, p(e_2 at t) is generally around y.

I take it that, in these circumstances, x would be very much higher than y. Given that there was no decay from atom III to atom IV, the chance of decay assessed just before t would be very much higher if there had been decay to atom II rather than not. If there had not been, then there would have been no decay to atom IV at all. I note that this seems akin to the *path-specific* solution that Dowe adopts (or perhaps Beebee's development of it) (Dowe, this volume: 34–6; Beebee, this volume: 45–9). The only point I wish to make here is that such cases do not pose a problem for my version of the counterfactual approach to causation. I will shortly consider Beebee's charge that, in effect, my approach has the counterintuitive upshot that any chance-lowering causal process will count as causal.

Doug Ehring's Case A involves two particles, α and β, colliding at t. The laws say that one and only one will be destroyed and each has a 50% chance of it being them. The survivor jumps noncontinuously to a space-time point e. The particles travel probabilistically rather than deterministically from point to point. Ehring claims that my approach cannot capture the fact that, in a particular case, it is β's location at t which is a cause of the location of a particle at e (α having been destroyed at t) (Ehring, this volume: 59–60, 67–70). The problem for my proposal is meant to be that the particle at e will Σ-probabilistically depend upon both α and β (with the other in Σ) so I must rely upon there being a missing intermediary (to wheel in clause (II)). But, from the description of the case, there is no missing non-actual intermediary.

The first point to make is that it is important to distinguish between two events at e: first, the occurrence of a particle at e; second the occurrence of the β-particle at e. It is clear that, if the focus of our interest was the latter event, it would Σ-probabilistically depend upon β at t and not upon α at t. If the event described as the occurrence of a particle at e was just the occurrence of the β-particle at e in the circumstances envisaged, then this point would constitute my answer to Ehring's problem case. α fails clause (I) of my account whereas β does not. However, I take it that he would insist that he did not have in mind the event of a particular particle, the β-particle, being at e. Rather he had in mind the less specific event of there just being a particle at e. He might claim that it is not essential to the event of there just being a particle at e that it is the β-particle. But if that is right, then there is a non-actual event that occurs when there is the α-particle at e which does not occur when there is the β-particle at e, namely that there is the α-particle at e. In which case, there is a non-actual event – there being an α-particle at e – upon which there being a particle at e would Σ-probabilistically depend if we placed β in Σ.

Ehring distinguishes his case from Schaffer's Trumping Cases by noting that there is an intrinsic difference in the processes involved in Case A whereas there is not in Schaffer's Trumping Cases. Some have viewed this to be an Achilles' heel of the Trumping Cases. Whatever the merits of that, the difference between Case A and the Trumping Cases has enabled me to avoid the challenge of Case A.

In fact, I do not think that Trumping Cases present any more difficulty for my theory. Let me briefly indicate why this is so. Here is a standard trumping case. At noon, Merlin casts the first spell of the day: to turn the prince into a frog at midnight. Later on, at 6 p.m., Morgana also casts a spell to turn the prince into a frog at midnight. At midnight, the prince turns into a frog. It is a law of magic that, if s is the first spell of the day and its aim is to bring about a certain result at midnight, then only this spell will have influence upon what happens at midnight. Hence, only the first spell, the claim runs, is a cause (Schaffer 2000a: 165). To establish that Merlin's spell is a cause, we may put Morgana's spell in Σ. In this case, Merlin's spell satisfies clauses (I) to (III) of my proposal. At first glance, it might seem as if the same would hold of Morgana's spell (if we put Merlin's spell in Σ). However, there is a difference. Putting Merlin's spell in Σ would (we may suppose) make Morgana's spell the first spell of the day with the aim of bringing about a certain result at midnight. The event of Morgana's being the first relevant spell of the day is non-actual. It is not in the case of Merlin's spell being the first relevant spell of the day. That means that the prince turning into a frog would be probabilistically Σ-dependent on a non-actual event (with Merlin's spell in Σ) in the case of Morgana. So Morgana's spell is not a cause.

Couldn't we fix up a version of the same problem for Merlin's spell? When we put Morgana's spell in Σ, we might say that Merlin's spell is an event of being the sole spell with the aim of bringing something about at midnight. I deny that there is such an event of being the sole spell in the circumstances described. But how can I deny that being the sole spell is an event when I insist that being the first spell of the day is an event? The answer lies in the structure of the case. I claim that a sufficient condition for something being an event is that its distinctive characterization figures in a relevant law of nature. That's precisely what Schaffer insists is the case in his trumping case. Even if I am wrong about this, the prince turning into a frog does not probabilistically Σ-depend upon the event of being the sole spell of the day, only the first. Thus, either way, we have the required asymmetry between Merlin's spell and Morgana's spell.[4]

Some of *Jonathan Schaffer's Overlapping Cases* raise much the same issue as Ehring's Case A. Here is one of the strongest examples. U238 and Ra226 are placed in a box at t_0. Each has a certain probability of decaying and producing an alpha particle. At t_1, there is an atom of Th234, an alpha particle, and Ra226. Hence we know that, although both Ra226 and U238 may decay producing an alpha particle, in this case it was U238. However, both are chance-raisers of the presence of the alpha particle. Suppose further that both decay spontaneously and immediately, and at least one of the atoms is in a superposition of location including positions at which the other atom is located. Then, Schaffer asserts, there will be no factors which could otherwise decide whether Ra226 or U238 decayed (Schaffer 2000b: 41–2).

Once more this overlooks the fact that two distinct alpha particles would be produced. Ra226 would not raise the chance of the distinct alpha particle produced by U238. If Schaffer invites us, as Ehring seemed to do, to focus on the event of an alpha particle (never mind which) being produced, then one of the more particular events will be the non-actual intermediary which the actually nondecaying atom's causal chain would involve in counterfactual circumstances. As a result, the alpha particle resulting from Ra226 in counterfactual circumstances with U238 in Σ would probabilistically Σ-depend upon a non-actual event.

Some cases of overlapping do not have the damaging feature just outlined by ascending to the world of magic. Suppose that Merlin casts a spell with 0.5 chance of turning the king and the prince into frogs and Morgana casts a spell with a 0.5 chance of turning the queen and the prince into frogs. As it happens, the king and the prince are turned into frogs. Schaffer argues that we may take it that this is a case of immediate causation and that Morgana's spell raised the chance of the prince turning into a frog but did not cause it (it was Merlin's spell that was responsible) (Schaffer 2000b: 40–1; see also Tooley, this volume: 90–1).

I don't deny that there was a point at which Morgana's spell raised the chance of the prince turning into a frog. However, I claim that at the time just before the prince turns into a frog, it does not. By then, the spell has proven ineffective. The fact that the queen is not turned into a frog demonstrates that this is the case. To argue that Morgana's spell raises the chance of the prince turning into a frog right up until the very time that the prince turns into a frog is to beg the question against the chance-raising proposal. It is not as if the description of the case requires that this claim be made. The intuitive plausibility of the thesis that causes raise the chances of effects presents reason for not making this claim.

It might be argued that I have not properly learnt the lessons of indeterminism. Am I not claiming that it is determinately not the case that Morgana's spell is a cause of the prince turning into a frog just before this happens? No – this is in the structure of the case. Schaffer is claiming that because the queen does not change into a frog, it is determinately not the case that Morgana's spell caused the prince to turn into a frog. All I am claiming is that if this is so, then it will show up in a failure of chance-raising just before the occurrence of the prince turning into a frog.

Suppose that Merlin's spell is put in the Σ-set. Will my proposal then have the verdict that Morgana's spell causes the prince to turn into a frog? If we imagine that very process in the context envisaged, then one might suppose that the answer is already obvious. Morgana's spell would not be a cause because it didn't raise the chance just before the prince turned into a frog. However, even though Merlin's and Morgana's spells are independent, it might be argued that, in the changed context, it still might be the case Morgana's spell was generally successful. It just didn't happen to be in the actual world. But even conceding this, there is a non-actual event upon which the prince turning into a frog will probabilistically Σ-depend: the queen turning into a frog. If the queen turned into a frog, the chance of the prince turning into a frog, assessed just before the prince turns into a frog, would be 0.5. Whereas, if the queen failed to turn into a frog, the chance of the prince turning into a frog, assessed just before the prince turns into a frog, would be 0 (again given the structure of the case). Hence Morgana's spell would fail clause (II) of my account.

In *Cartwright's Weed Case*, a weed in a garden is sprayed with a defoliant. This decreases the chance it will survive from 0.7 to 0.3. The plant is sick for six months but then recovers. There is a causal process from spraying defoliant on the leaves to the recovery. For instance, in counterfactual terms, if the defoliant had not been sprayed on the leaves, the plant would not have lost all its leaves. If the plant had not lost all its leaves, it would not have sprouted a whole host of new ones. If the plant had not sprouted a whole host of new ones, then it would not have been healthy six months later. Yet, in spite of this, it seems clear that spraying defoliant on the leaves was not a cause of the plant's health six months later (see Cartwright 1983: 28, in part a reprint of Cartwright 1979).

At first glance, this case seems to present no problem at all for my approach. The characterization of the causal process just given appealed to successive counterfactuals. However, my approach does not take the ancestral of counterfactuals as distinctive of causation. Focusing on my preferred chance-raising counterfactual, it seems that spraying defoliant on the leaves would not, in general, raise the probability of health in the plant significantly over circumstances in which the defoliant is not sprayed on the leaves.

Unfortunately, as Beebee and Hitchcock point out, matters are not quite so straightforward for accounts like mine (Beebee, this volume: 48–9; Hitchcock, this volume: 146–9). Suppose we put in Σ the negative event of not getting deadly leaf disease. In which case, we would have to focus on all the worlds in which the plant *did* get deadly leaf disease. It would seem that, in these worlds, spraying defoliant on the plant would raise the chance of the plant's survival because the infection would be extremely limited. So, it is argued, I have to concede that my account would yield the verdict that spraying defoliant was a cause after all.

I do not think so. In this case, the plant's survival would probabilistically Σ-depend upon non-actual events. One example would be deadly leaf infection being stopped from spreading. I can imagine that some might question whether this is a genuine event. To the extent that it is legitimate to raise the question here, it is also legitimate to raise the question over the putative negative event of not getting deadly leaf disease. My point is just that if one is going to be liberal about what counts as events in this case, then the same, indeed a lesser, liberality will save my theory from counterexample.

5 Conclusion

I have defended my preferred counterfactual approach to causation against challenges drawn from the apparent circularity attendant on appealing to counterfactuals, the nontransitivity of causation, and various counterexamples to theories based on chance-raising. It seems to me that the result of the discussion is that the prospects of a counterfactual theory of causation are good, contrary to the claims of some recent critics. Nevertheless, there are still some areas of concern. If I were going to single out one, it would be the question of whether counterfactual approaches can capture a defensible notion of causal asymmetry. I am not convinced that there is a genuine causal asymmetry over and above the kind of macro-asymmetries that present little difficulty for the counterfactual theorist. But if there were, it seems to me that the prospects for a reductive analysis of causation in terms of counterfactuals would be attenuated. That is not to say that appeal to counterfactuals would not lie right at the heart of the proper characterization of causation even in those disappointing circumstances.

Notes

1 I have simplified clause (II) and deleted clause (IV), the latter is motivated by cases of catalysts and anticatalysts (see Noordhof 1999: 115–20). For further discussion of clause (II) and its proper formulation, see Sungho Choi (2002) and Noordhof (2002).

2 Ramachandran flirts with this line but concedes it probably won't deal with all cases (1998: 469–70). In his later work, he abandons all hope (see Ramachandran, this volume).

3 My appeal to raising the chance of an event at a time is to deal with cases of late preemption. For details, see Noordhof (1999). For more on hasteners and delayers, see Penelope Mackie (1992), who got me thinking about these issues.

4 In my paper 'In Defence of Influence' I noted that clause (IV) of my full theory would also deal with trumping cases (see Noordhof 2001: 323, fn. 1). However, I thought it worthwhile to note here that it is not clear that it is needed for this purpose. In any event, its rationale was derived from cases of catalysts and anticatalysts (see Noordhof 1999: 115–20). I was partly inspired to make my present defence against trumping through reading Stephen Barker's chapter in this volume. However, there is an important difference. If you appeal indiscriminately to events, states, conditions or anything else nonactual, you will be able to fix up problematic dependencies to discredit genuine causes.

Bibliography

Adams, E. (1975) *The Logic of Conditionals*, Dordrecht: Reidel.

—— (1993) 'On the Rightness of Certain Counterfactuals', *Pacific Philosophical Quarterly* 74: 1–10.

—— (1998) 'Remarks on Wishes and Counterfactuals', *Pacific Philosophical Quarterly* 79: 191–5.

Anderson, A. R. (1951) 'A Note on Subjunctive and Counterfactual Conditionals', *Analysis* 12: 35–8.

Anscombe, G. E. M. (1971) *Causality and Determination*, Cambridge: Cambridge University Press.

Armstrong, D. M. (1968) *A Materialist Theory of the Mind*, London: Routledge and Kegan Paul.

—— (1983) *What Is a Law of Nature?*, Cambridge: Cambridge University Press.

—— (1997) *A World of States of Affairs*, Cambridge: Cambridge University Press.

—— (1999) 'The Open Door', in H. Sankey (ed.) *Causation and Laws of Nature*, Dordrecht: Kluwer, pp. 175–85.

Armstrong, D. M. and Heathcote, A. (1991) 'Causes and Laws', *Noûs* 25: 63–73.

Aronson, J. (1971) 'On the Grammar of "Cause"', *Synthese* 62: 249–57.

Barker, S. (1998) 'Predetermination and Tense Probabilism', *Analysis* 58: 290–6.

—— (1999) 'Counterfactuals, Probabilistic Counterfactuals and Causation', *Mind* 108: 427–69.

—— (2003a) 'Counterfactual Analysis of Causation: The Problem of Effects and Epiphenomena Revisited', *Noûs* 37: 133–50.

—— (2003b) 'A Dilemma for the Counterfactual Analysis of Causation', *Australasian Journal of Philosophy* 81: 62–77.

Beebee, H. (1997) 'Taking Hindrance Seriously', *Philosophical Studies* 88: 59–79.

—— (forthcoming), 'Causing and Nothingness', in J. Collins, E. J. Hall and L. A. Paul (eds) *Counterfactuals and Causation*, Cambridge MA: MIT Press.

Bennett, J. (1974) 'Counterfactuals and Possible Worlds', *Canadian Journal of Philosophy* 4: 381–402.

—— (1987) 'Event Causation: The Counterfactual Analysis', in J. E. Tomberlin (ed.) *Philosophical Perspectives I*, Atascadero CA: Ridgeview, pp. 367–86.

Braithwaite, R. B. (1953) *Scientific Explanation*, Cambridge: Cambridge University Press.

Carroll, J. W. (1994) *Laws of Nature*, Cambridge: Cambridge University Press.

Cartwright, N. (1979) 'Causal Laws and Effective Strategies', *Noûs* 13: 419–38. Reprinted in N. Cartwright (1983), *How the Laws of Physics Lie*, Oxford: Oxford University Press.

—— (1983) *How the Laws of Physics Lie*, Oxford: Oxford University Press.
—— (1989) *Nature's Capacities and their Measurement*, Oxford: Clarendon Press.
Choi, S. (2002) 'The "actual events" clause in Noordhoff's account of Causation', *Analysis* 62/1: 41–6.
Collins, J. (2000) 'Preemptive Prevention', *Journal of Philosophy* 97: 223–34.
Dowe, P. (1992) 'Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory', *Philosophy of Science* 59: 195–216.
—— (1999) 'The Conserved Quantity Theory of Causation and Chance Raising', *Philosophy of Science* 66 (Proceedings): S486–S501.
—— (2000a) 'Causing, Promoting, Preventing, Hindering', in M. Ledwig, W. Spohn and M. Esfeld (eds), *Current Issues in Causation*, Parderborn: Mentis-Verlag, pp. 69–84.
—— (2000b) *Physical Causation*, Cambridge and New York: Cambridge University Press.
—— (2001) 'A Counterfactual Theory of Prevention and "Causation" by Omission', *Australasian Journal of Philosophy*, 79/2 (June): 216–26.
—— (Unpublished) 'Is Causation Influence?'
Dretske, F. I. (1977) 'Laws of Nature', *Philosophy of Science* 44: 248–68.
Edgington, D. (1995) 'On Conditionals', *Mind* 104: 235–329.
Eells, E. (1991) *Probabilistic Causality*, Cambridge: Cambridge University Press.
Ehring, D. (1997) *Causation and Persistence*, New York: Oxford University Press.
—— (forthcoming) 'Part-Whole Physicalism and Mental Causation', *Synthese*.
Elwood, J. M. (1992) *Causal Relationships in Medicine*, New York: Oxford University Press.
Fine, K. (1975) 'Critical Notice – Counterfactuals', *Mind* 84: 451–8.
Ganeri, J., Noordhof, P. and Ramachandran, M. (1996) 'Counterfactuals and Preemptive Causation', *Analysis* 56: 216–25.
—— (1998) 'For A (Revised) PCA Analysis', *Analysis* 58: 45–7.
Glassner, B. (1999) *The Culture of Fear*, New York: Basic Books.
Goldman, A. (1967) 'A Causal Theory of Knowing', *Journal of Philosophy* 64, 12: 355–72.
Good, I. J. (1961) 'A Causal Calculus I', *British Journal for the Philosophy of Science* 11: 305–18.
—— (1962) 'A Causal Calculus II', *British Journal for the Philosophy of Science* 12: 43–51.
Goodman, N. (1955) *Fact, Fiction and Forecast*, Indianapolis: Bobbs-Merrill.
Grünbaum, A. (1973) *Philosophical Problems of Space and Time*, Dordrecht:Reidel.
Hall, E. J. (2000) 'Causation and the Price of Transitivity', *Journal of Philosophy* 97: 198–223.
—— (forthcoming) 'Two Concepts of Causation', in J. Collins, N. Hall and L. Paul (eds) *Counterfactuals and Causation*, Cambridge MA: MIT Press.
Hausman, D. M. (1998) *Causal Asymmetries*, Cambridge: Cambridge University Press.
Hesslow, G. (1976) 'Two Notes on the Probabilistic Approach to Causation', *Philosophy of Science* 43: 290–2.
Hitchcock, C. (1995a) 'The Mishap at Reichenbach Fall: Singular vs. General Causation', *Philosophical Studies* 78: 257–91.
—— (1995b) 'Discussion: Salmon on Explanatory Relevance', *Philosophy of Science* 62: 304–20.
—— (1996) 'The Mechanist and the Snail', *Philosophical Studies* 84: 91–105.
—— (2001a) 'The Intransitivity of Causation Revealed in Equations and Graphs', *Journal of Philosophy* 98: 273–99.
—— (2001b) 'A Tale of Two Effects', *The Philosophical Review* 110: 361–96.
—— (2003) 'Of Humean Bondage', *British Journal for the Philosophy of Science* 54: 1–25.

—— (forthcoming) 'Do All and Only Causes Raise the Probabilities of Effects?', in J. Collins, N. Hall and L. Paul (eds) *Causation and Counterfactuals*, Cambridge MA: MIT Press.

Hume, D. (1739–40) *A Treatise of Human Nature*, London.

—— (1748) *An Enquiry Concerning Human Understanding*, London.

Humphreys, P. (1981) 'Probabilistic Causality and Multiple Causation', in D. A. Peter and R. N. Giere (eds) *PSA 1980*, East Lansing, Michigan, pp. 25–37.

Jackson, F. (1977) 'A Causal Theory of Counterfactuals', *Australasian Journal of Philosophy* 55: 3–21.

Johnson, D. (1991) 'Induction and Modality', *Philosophical Review* 100: 399–430.

Kim, J. (1973) 'Causes and Counterfactuals', *Journal of Philosophy*, 70: 570–2. Reprinted in Ernest Sosa and Michael Tooley (eds) (1993), *Causation*, Oxford: Oxford University Press, pp. 205–7.

Kvart, I. (1986) *A Theory of Counterfactuals*, Indianapolis: Hackett Publishing Company.

—— (1991a) 'Counterfactuals and Causal Relevance', *Pacific Philosophical Quarterly*, 314–37.

—— (1991b) 'Transitivity and Preemption of Causal Impact', *Philosophical Studies* 64: 125–60.

—— (1994) 'Overall Positive Causal Impact', *Canadian Journal of Philosophy*, 24: 205–8.

—— (1997) 'Cause and Some Positive Causal Impact', *Philosophical Perspectives*, 11: *Mind, Causation and World*, J. Tomberlin (ed.): 401–32.

—— (2001a) 'Causal Relevance', in Bryson Brown and John Woods (eds) *New Studies in Exact Philosophy: Logic, Mathematics and Science* (selected contributions to the *Exact Philosophy* Conference, May 1999), Vol. II, Hermes Scientific Publications, pp. 59–90.

—— (2001b) 'A Counterfactual Theory of Cause', *Synthese* 127/3: 389–427.

—— (2001c) 'Lewis' "Causation as Influence"', *Australasian Journal of Philosophy* 79/3:411–23.

—— (2002) 'Probabilistic Cause and the Thirsty Traveler', *Journal of Philosophical Logic* 31/2: 139–79.

—— (2003) 'Causation: Probabilistic and Counterfactual Analyses', in J. Collins, N. Hall and L. Paul (eds) *Causes and Counterfactuals*, Cambridge MA: MIT Press.

—— (forthcoming a) 'Probabilistic Causation and Mental Causation'.

—— (forthcoming b) 'Partial Cause Neutralizers'.

—— (forthcoming c) 'Cause, Time and Manner'.

Langton, R. and Lewis, D. K. (1998) 'Defining Intrinsic', *Philosophy and Phenomenological Research* 58: 333–45.

Lewis, D. K. (1970) 'How to Define Theoretical Terms', *Journal of Philosophy* 67: 427–46.

—— (1973a) *Counterfactuals*, Cambridge MA: Harvard University Press. Also Oxford: Basil Blackwell.

—— (1973b) 'Causation', *Journal of Philosophy* 70: 556–67. Reprinted, with postscripts, in *Philosophical Papers, Vol. 2.*

—— (1979) 'Counterfactual Dependence and Time's Arrow', in *Noûs* 13: 455–76. Reprinted, with postscripts, in *Philosophical Papers*, *Vol. 2*, New York: Oxford University Press, pp. 32–66. Page references to this volume.

—— (1986) *Philosophical Papers*, *Vol. 2*, Oxford: Oxford University Press.

—— (2000) 'Causation as Influence', *Journal of Philosophy*, 97/4: 182–98.

Mackie, P. (1992) 'Causing, Delaying and Hastening: Do Rains Cause Fires?' *Mind* 101/403: 483–500.

Madden, E. H. and Harré, Rom (1975) *Causal Powers*, Oxford: Blackwell.

Martin, C. B. (1993) 'Power for Realists', in John Bacon, Keith Campbell and Lloyd Reinhardt (eds), *Ontology, Causality and Mind*, Cambridge: Cambridge University Press.

Martin, R. M. (1966) 'On Theoretical Constructs and Ramsey Constants', in *Philosophy of Science* 33: 1–13.

McDermott, M. (1995) 'Redundant Causation', in *The British Journal for the Philosophy of Science* 40: 523–44.

—— (1997) 'Metaphysics and Conceptual Analysis: Lewis on Indeterministic Causation', *Australasian Journal of Philosophy* 75: 396–403.

—— (2002) 'Causation: Influence Versus Sufficiency', *The Journal of Philosophy* 97: 84–101.

Mellor, D. H. (1986) 'Fixed Past, Unfixed Future', in Barry Taylor (ed.) *Contributions to Philosophy: Michael Dummett*, The Hague: Nijhoff, pp. 166–86.

—— (1995) *The Facts of Causation*, London: Routledge.

Menzies, P. (1989) 'Probabilistic Causation and Causal Processes: A Critique of Lewis', *Philosophy of Science* 56: 642–63.

—— (1996) 'Probabilistic Causation and the Pre-emption Problem', *Mind* 105: 85–117.

—— (1998) 'How Justified are Humean Doubts about Intrinsic Causal Relations?', in *Communication and Cognition* 31: 339–64.

—— (1999) ' Intrinsic versus Extrinsic Conceptions of Causation', in H. Sankey (ed.) *Causation and Laws of Nature*, Dordrecht: Kluwer, pp. 313–30.

Noordhof, P. (1998a) 'Problems for the M-Set Analysis of Causation', *Mind* 107/426: 457–63.

—— (1998b) 'Critical Notice: Causation, Probability and Chance, in D.H. Mellor, *The Facts of Causation*', *Mind* 107/428: 855–77.

—— (1999) 'Probabilistic Causation, Preemption and Counterfactuals', *Mind* 108/429: 95–125.

—— (2000) 'Ramachandran's Four Counterexamples', *Mind* 109: 315–24.

—— (2001) 'In Defence of Influence?', *Analysis* 61/4: 323–7.

—— (2002) 'Sungho Choi and the "actual events" Clause', *Analysis* 62/1: 46–7.

Paul, L. A. (1998a) 'Problems with Late Pre-emption', *Analysis* 58: 48–54.

—— (1998b) 'Keeping Track of the Time: Emending the Counterfactual Analysis of Causation', *Analysis* 58: 191–8.

Pearl, J. (2000) *Causality: Models, Reasoning and Inference*, Cambridge: Cambridge University Press.

Popper, K. (1956) 'The Arrow of Time', *Nature* 177: 538.

Ramachandran, M. (1997) 'A Counterfactual Analysis of Causation', *Mind* 106: 263–77.

—— (1998) 'The M-Set Analysis of Causation: Objections and Responses', *Mind* 107: 465–71.

—— (2000) 'Noordhof on Probabilistic Causation', *Mind* 109/434: 309–13.

—— (forthcoming) 'A Counterfactual Analysis of Indeterministic Causation', in J. Collins, N. Hall and L. A. Paul (eds) *Causation and Counterfactuals*, Cambridge MA: MIT Press.

Ramsey, F. P. (1929) 'Theories', first published in R. B. Braithwaite (ed.) *The Foundations of Mathematics*, London: Routledge and Kegan Paul (1931), chapter 9.

Reichenbach, H. (1956) *The Direction of Time*, Berkeley and Los Angeles: University of California Press.

Rosen, D. (1978) 'Discussion: In Defence of a Probabilistic Theory of Causality', *Philosophy of Science* 45: 604–13.

Salmon, W. (1978) 'Why Ask "Why?"?', *Proceedings and Addresses of the American Philosophical Association* 51/6: 683–705.

—— (1980) 'Probabilistic Causality', *Pacific Philosophical Quarterly* 61: 50–74.
—— (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.
—— (1997) 'Causality and Explanation: A Reply to Two Critiques', *Philosophy of Science* 64: 461–77.
—— (1998) *Causality and Explanation*, New York: Oxford University Press.
Schaffer, J. (2000a) 'Trumping Preemption', *Journal of Philosophy* 97/4: 165–81.
—— (2000b) 'Overlappings: Probability – Raising without Causation', *Australasian Journal of Philosophy* 78: 40–6.
—— (2000c) 'Causation by Disconnection', *Philosophy of Science* 67: 285–300.
—— (2001a) 'Causes as Probability Raisers of Processes', *Journal of Philosophy* 98: 75–92.
—— (2001b) 'Causation, Influence, and Effluence', *Analysis* 61: 11–19.
Scriven, M. (1956–57) 'Randomness and the Causal Order', *Analysis* 17: 5–9.
Slote, M. (1978) 'Time in Counterfactuals', *Philosophical Review* 87: 3–27.
Sosa, E. and Tooley, M. (eds) (1993) *Causation*, Oxford: Oxford University Press.
Spirtes, P., Glymour, C. and Scheines, R. (1993) *Causation, Prediction, and Search*, New York: Springer-Verlag. (Second edition Cambridge MA: MIT Press, 2000.)
Stalnaker, R. C. (1968) 'A Theory of Conditionals', in Nicholas Rescher (ed.) *Studies in Logical Theory*, Oxford: Blackwell, and reprinted in Ernest Sosa (ed.) (1975) *Causation and Conditionals*, Oxford: Oxford University Press, 165–79.
Strawson, G. (1989) *The Secret Connexion: Causation, Realism, and David Hume*, Oxford: Oxford University Press.
Suppes, P. (1970) *A Probabilistic Theory of Causality*, Amsterdam: North-Holland Publishing Company.
—— (1984) *Probabilistic Metaphysics*, Oxford: Blackwell.
Tichy, P. (1976) 'A Counterexample to the Stalnaker–Lewis Analysis of Counterfactuals', *Philosophical Studies* 29: 271–3.
Tooley, M. (1977) 'The Nature of Laws', *Canadian Journal of Philosophy* 7: 667–98.
—— (1984) 'Laws and Causal Relations', in P. A. French, T. E. Uehling and H. K. Wettstein (eds) *Midwest Studies in Philosophy* 9, Minneapolis: University of Minnesota Press, pp. 93–112.
—— (1987) *Causation: A Realist Approach*, Oxford: Oxford University Press.
—— (1990a) 'The Nature of Causation: A Singularist Account', in David Copp (ed.) *Canadian Philosophers*, *Canadian Journal of Philosophy*, Supplement 16: 271–322. Reprinted in Jaegwon Kim and Ernest Sosa (eds) (1999) *Metaphysics: An Anthology*, Blackwell: Oxford, pp. 458–82.
—— (1990b) 'Causation: Reductionism Versus Realism', *Philosophy and Phenomenological Research* 50: 215–36.
—— (2003) 'The Stalnaker-Lewis Approach to Counterfactuals', *Journal of Philosophy* 100: 371–7
van Fraassen, B. C. (1989) *Laws and Symmetry*, Oxford: Clarendon Press.
Woodward, J. (1990) 'Supervenience and Singular Causal Claims', in D. Knowles (ed.) *Explanation and its Limits*, Cambridge: Cambridge University Press, 211–46.
Yablo, S. (2002) 'De Facto Dependence', *Journal of Philosophy* 99: 130–48.

Index