JOHN BIGELOW AND ROBERT PARGETTER

METAPHYSICS OF CAUSATION

Erkenntnis, 1990 (Vol. 33, No. 1, 89-119).

I have slightly changed the text in a few places, to explain terminology and improve clarity.

1. CAUSATION AND NECESSARY CONNECTION

There is a widely-held and plausible, yet mistaken view which links causation with necessity. Causation is construed as some sort of 'necessary connection'. This connection is conceived as entailing either that the cause is a *necessary condition* for the effect, or else that given the cause, the effect was a *necessary consequence* of that cause. That is to say, the cause is conceived as either a necessary, or else a sufficient condition for the effect – or else it is both a necessary and a sufficient condition for the effect.[1]

Some recent theorists have preserved this general structure behind their theory, but have weakened the notion of 'necessary and sufficient conditions'. Instead of the traditional construal in terms of *impossibility* of finding cause without effect, or effect without cause, it has been suggested that we should content ourselves with some sort of (relative) *improbability* of finding cause without effect, or effect without cause. Such theories are called probabilistic accounts of causation, but they are close enough in spirit to a necessitarian account of causation to be included, for our purposes, together with 'necessary connection' theories, since they preserve the idea that causation is some kind of *modal* relation.[2]

---

[1] For a classic collection of important papers, and further references, see E. Sosa, *Causation and Conditionals*, Oxford University Press, London, 1975.

[2] We have in mind particularly the defence of probabilistic theories by D. K. Lewis in 'Counterfactual Dependence and Times Arrow', reprinted with notes in his *Philosophical Papers II*, Oxford University Press, London, 1986. There are, however, dozens of supporters of such theories. One key early account is that of P. Suppes, *A Probabilistic Theory of Causality*, North-Holland, Amsterdam, 1970.

[N.B. The chief 'modal' relations are necessity and possibility. Probability is a degree of necessity. For example, if B is a necessary consequence of A then the probability that B occurs, in cases where A occurs, is equal to 1. Thus necessity is the maximum degree of probability. Lower probabilities, lying between 0 and 1, represent lower degrees of necessity. Thus, the authors later refer to probability as a 'weakened form' of necessity.]

In the initial sections of this paper, we present arguments in support of the growing number of philosophers who say that all such theories are misguided. Rejection of necessitarian (i.e. modal) theories will leave us in urgent need of a substitute. In the later parts of this paper [not included here], we support and extend a rival approach to causation, one which grounds causation more intimately in the details of scientific theories.

Of course a 'necessary connection' theory of causation owes us some account of what kind of 'necessity' it rests upon. Either one of the realist or one of the non-realist theories of necessity may be added to some necessitarian (or probabilistic) account of causation, to give the full theory. But regardless of how the theory of necessity (or probability) is spelt out, we believe both necessitarian and probabilistic theories of causation should be rejected. A cause may be neither a necessary nor a sufficient condition for an effect. The effect could have come about without the cause, either from some other cause or by no cause at all, and consequently the cause is not a necessary condition for the effect. Nor can the occurrence of an event always be taken to ensure a high probability that it was preceded by a certain cause, nor even to have increased the probability of that cause above what it would otherwise have been. Neither necessity nor its probabilistic weakening are essential to causes, as we shall argue more fully below. Similarly, neither sufficiency nor its probabilistic weakening are essential to causes. A cause need not be a sufficient condition for the effect, and may not even ensure an increased probability for that effect.

We have learned this largely from Hume.[3] Hume's contributions to the theory of causation have a theological background. French theologians,

[3] David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, Oxford University Press, London, 1888.

notably Descartes and Malebranche, belonged to a theological tradition which insisted that God could not be fettered by any constraints whatever upon His freedom. Hence given a cause, any cause, God cannot be thereby compelled to permit the effect to follow. If cause is followed by effect, this can only be by the grace of God, by an entirely free choice, on the part of God, to permit the effect to follow. God could intervene and present us with a miracle whenever He chooses. Hence the cause is not, by itself, a logically sufficient condition for the effect. It is only the cause together with the will of God which yields a sufficient condition for the effect. Given just the cause alone, at any time prior to the effect, it is possible for God to choose not to permit the effect. Hence it is possible for the cause to occur and the effect not to follow. That is to say, the cause is not a sufficient condition for the effect. Nor can theologians like Descartes allow that the cause is necessary for the effect. God, being omnipotent, could have brought about the very same effect simply by willing it, or by the mediation of some other quite different cause.

Take the conclusion of this theological argument, then remove God from the scene, and the result is Hume's theory of causation. Instead of asking us to admit that God could choose not to permit the expected effect to follow a given cause, Hume asks us to admit simply that the effect could fail to follow a given cause. And he is right. (In fact he asks us to imagine the effect failing to follow, and he takes imaginability as a guide to possibility. The shift from a theological to a psychological argument is not an unqualified improvement in the strength of the argument. Yet the conclusion is a compelling one, however doubtful the route which brought us there.) Hume is right in stressing that the effect could fail to follow – and this is true not only in the sense of logical possibility, but of empirical possibility as well. Causes are not sufficient conditions. And the same applies to the claim that causes are necessary conditions. Sometimes the effect would have or could have occurred even if the cause had been absent.

The denial that a cause is a sufficient condition for its effect leads Hume to look elsewhere for his account of causation. Hume's attention is drawn in two directions: one 'outwards', the other 'inwards'. He refers us to the external facts about *regularities* in nature; and he refers us to the internal *expectations* that arise in us after exposure to such regularities.

In one sense, then, Hume denies that a cause is a sufficient condition for its effect. He denies the 'sufficiency' of the cause, in the sense in which 'sufficiency' involves some genuine modality. That is to say, he denies the sufficiency of the cause, if 'sufficiency' is taken under a *realist* construal. Yet he does not deny the sufficiency of the cause, if 'sufficiency' is taken, as we might say, more subjectively. He does not deny that the cause is a sufficient condition for its effect *in the sense that* such effects do always follow and we would be surprised if any given one of them didn't. For the Humean, this is how laws of nature are to be distinguished from 'mere' regularities, not by further information about how the law relates to things in those parts of nature which are being described, but by information about how the law relates to people, their other opinions, their purposes, habits, expectations and so forth.

It is unnecessary here to offer a critique of Hume's account (and Humean accounts generally).[4] We do not accept a non-realist theory of laws. But we do accept the lesson Hume taught us that causes are not sufficient for their effects (if 'sufficiency' is underpinned by more than just regularity and subjective expectation). Indeed, we go much further: we also deny the prior assumption that a cause is always necessary or sufficient for its effect, even in the minimal sense that it is an instance of a regularity which has some non-trivial status. Hume did go too far in his rejection of necessity in laws of nature, but he did not go far enough in his rejection of the necessitarian account of causation.

2. CAUSES ARE NOT NECESSARY CONDITIONS

We will argue that a cause is neither a necessary nor a sufficient condition for its effect, setting aside Humean and other primarily subjective senses of necessary and sufficient conditions. One event may cause another, and yet fail to be a necessary condition for that other event because there is a 'fail safe' backup system which would have brought about the same effect if the actually operative system had failed.

[4] For a good recent survey and attack on Humean theories of laws (and indirectly, on cause), see D. M. Armstrong, *What is a Law of Nature?*, Cambridge University Press, Cambridge, 1983.

Consider for instance the food that nourishes you. Eating the particular slice of bread that you did eat will cause a variety of effects; but eating that specific slice was not a necessary condition for the production of those effects. If you hadn't eaten that slice, you could have eaten another.

You might suspect that eating a different slice would have had slightly different effects. Yet there is no guarantee of this. It is quite conceivable that a wide variety of food intakes could have produced exactly the same outcome. Living things involve a variety of homeostatic systems which aim to preserve a constant state despite varying causal inputs. In general, admittedly, they do fail to maintain absolute constancy – or so it is natural for us to speculate. However, this omnipresence of imperfection, if it exists, is a contingent factor. There is nothing intrinsic to causation itself which entails that homeostatic systems must always be imprecise and imperfect. Indeed, the quantization of small-scale phenomena in physics would suggest that at least for some small-scale events, different causes could have precisely the same effect.

Lewis allows that a cause may not be a necessary condition for its effect. Yet he does explain causation, less directly, in terms of necessary connections. Lewis defends a theory which analyzes causation in terms of *chains* of necessary conditions. One event is a cause of another, he says, provided there is a chain of distinct (non-overlapping) events, beginning with the former and ending with the latter event, in which each of the events in the chain is a necessary condition for the one which follows. And for one event P to be a necessary condition for the following event Q, is for a specific counterfactual to hold, namely: that *if P had not occurred then Q would not have occurred either*. This counterfactual conditional is often written $\neg P \mathrel{\square\!\!\rightarrow} \neg Q$. In indeterministic cases, Lewis weakens this construal of necessary connection, replacing it by a probabilistic one. We will leave this aside, however, for the moment.[5]

While Lewis uses counterfactuals in his analysis, Mackie achieves the same kind of analysis using strict conditionals, i.e. conditionals of the

[5] D. K. Lewis, 'Causation', in Sosa, *Causation and Conditionals*, pp. 180-191; and reprinted with additional notes in D. K. Lewis, *Philosophical Papers II*. Lewis discusses indeterministic cases in 'Counterfactual Dependence and Time's Arrow'.

form “Necessarily, if P occurs then Q occurs as well.”, which can be symbolically represented as $\Box(P \rightarrow Q)$. If C is the conjunction of background conditions which determine whether the presence or absence of the cause *c* determines the presence or absence of the effect *e*, then the Lewis counterfactuals

c does occur $\Box\!\!\rightarrow$ *e* does occur
c does not occur $\Box\!\!\rightarrow$ *e* does not occur

are replaced by the strict conditionals

$\Box$(C holds and *c* does occur $\rightarrow$ *e* does occur)
$\Box$(C holds and *c* does not occur $\rightarrow$ *e* does not occur).[6]

Hence we derive a view summed up by Mackie’s mnemonic that a cause is an INUS condition for its effect, an *in*sufficient but necessary part of an unnecessary but sufficient condition.[7]

Lewis and Mackie agree that a cause is a ‘necessary condition’ for its effect; they differ only over how this is to be analyzed. Lewis appeals to counterfactuals, whose semantics draw upon indefinitely many features of the actual world. Mackie appeals to strict conditionals which include complex antecedents which in practice we can seldom state explicitly, yet which would yield a full account of the causal process if we could state them. In two different ways, then, Lewis and Mackie provide a background against which a cause is claimed to be a necessary condition for its effect.

These theories of causation have a number of merits. In particular, they do allow space for a number of different kinds of backup systems and homeostatic mechanisms. A cause may have a variety of remote effects in virtue of a chain of intermediate causes and effects. It may then happen that such a cause is not a necessary condition for its remote effects. If the cause had not occurred, then its chain would not have

[6] For pioneering work on counterfactuals and strict conditionals, see Nelson Goodman, *Fact, Fiction and Forecast*, Bobbs-Merrill, New York, 1965.
[7] J. L. Mackie, ‘Causes and Conditions’, in Sosa, *Causation and Conditionals*, pp. 15-48.

begun, and yet some other causal chain might have brought about the same final event. In such a case, the cause fails to be a necessary condition in the counterfactual sense. It is not true that if the cause had not occurred, then the effect would not have occurred. Yet Lewis does count it as a cause, because it contributes to a causal chain, a chain of necessary conditions. Mackie, too, can count it as a cause, because it is a necessary part of some sufficient condition – in this case, the sufficient condition which embraces the whole chain of necessary conditions described by Lewis plus the absence of back-up chains.

Despite the merits of theories like those of Lewis or Mackie, there are also some drawbacks. These drawbacks pertain not just to the alleged necessity of causes for effects, but also to any alleged sufficiency of causes for effects. Hence we will not treat necessary conditions and sufficient conditions separately, but will class them together as subspecies of the same basic theory that causation is to be analyzed in terms of a modal connection between cause and effect. It is that basic theory that we will be rejecting.

3. NEITHER NECESSARY NOR SUFFICIENT CONDITIONS

Aside from more or less detailed worries, there is a very general objection to theories which explain causation in terms of necessary or sufficient conditions for the effect. Such theories are too closely bound up with the assumption of some sort of *determinism* in nature.

By determinism, we do not mean simply the doctrine that every event has a cause. Even if we grant that every event has a cause, it does not follow that every event is 'determined' by the cause, unless the cause is taken to be a sufficient condition for the effect. Yet consider what we are committed to, if we take every event to have a cause, and we take causes to be sufficient conditions for their effects. It follows that without sufficient conditions, there will be no causes. That is to say, insofar as there is indeterminism in nature, to that extent we would need to abandon causation.

Yet despite the pervasiveness of indeterminacy in the subatomic realm, we have not been forced to abandon causal talk. In a variety of cases, it is

quite clear that we have caused some event to occur, even though it was possible that the event should have failed to occur. For instance, suppose an electron strikes a copper target causing an electron to dislodge from the inner shell of electrons of a copper atom. This in turn produces the emission of an X-ray photon, but the photon produced can be of various frequencies depending on from which of the other shells of the atom the electron comes which fills the space in the inner shell. The existence of more than one possible outcome does not affect the appropriateness of our saying that the bombarding electron, or creation of a space in the inner electron shell, *causes* the outcome which *does* eventuate.

Defenders of the idea that 'causes are conditions' may modify their theory to make room for causation under indeterminism. Instead of saying a cause is a 'sufficient' condition for an effect, we could say that the cause makes the effect 'very probable'. This would be the simplest replacement we could make for the notion of sufficient conditions. And yet, a variety of examples put pressure on us to by-pass that theory in favour of a less direct account. For instance a person having syphillis can develop paresis, and that case of paresis is caused by the syphillis; and yet the probability of developing paresis, even among those with syphillis, is small. We need a somewhat less direct account of the link between causation and probabilities. The most plausible candidate is the theory that a cause makes the effect more probable than it would have been in the absence of that cause.

This theory has a number of merits. It neatly fuses the idea of cause as a sufficient condition with that of cause as a necessary condition. In the *presence* of the cause, the effect is *more* probable (a weakened notion of a 'sufficient condition'); in the *absence* of the cause, the effect is *less* probable (a weakened form of a 'necessary condition').

This probabilistic account of causation gives extremely plausible accounts of a wide range of cases. Nevertheless, we argue that it is on the wrong track. Our reasons for resisting such a theory are of several sorts.

We believe that causation is a *local* feature of a cause-effect pair. What makes one thing a cause of another is entirely a matter of the nature of the cause, of the effect, and of what transaction occurs between them. Causation is, roughly speaking, a two-place relation, not an indefinitely-

many-place relation. We can leave the causal relation unaltered, even if we vary the context in which it occurs.

Thus, for instance, we may leave the causal relation unaffected, even if we institute a 'back-up system' which would have come into play had the cause failed to occur. Cases of this sort are familiar in the literature. Imagine for instance Gorbachev pressing the button that launches the rocket 'Glaznost', on its journey to Mars. The causal relation between his pressing the button and the launch occurring is exactly the same, whether or not there is a fail-safe mechanism which would over-ride the intended causal path, were Gorbachev to bungle. The relation between cause and effect is, we claim, independent of the presence of back-up systems waiting in the wings.

Causation is, we claim, a 'local' matter of the actual, physical transaction between cause and effect. Necessary and sufficient conditions, in contrast, are much more 'global' concerns. By instituting a back-up system, what was a necessary condition ceases to be a necessary condition. Yet the intrinsic character of the causal process is not altered.

It is important to clarify our claim here. It has to be acknowledged that there are such things as causal laws, and of course causal laws are, in our terms, 'global': the truth of a causal law depends on the character of a world as a whole, and not just on one of its constituents. But the truth of causal laws supervenes, we claim, on the existence of a pattern of causal transactions in the world. The law is (or entails) a *generalization* over causal transactions. (In fact we take a law to be something stronger than a generalization, as we believe laws involve some kind of modality, but that is beside the point here.) The transactions do not count as causal because they are subsumed under laws. The connection works the other way around: the laws hold because of the presence of local causal connections. Necessary and sufficient conditions generally are underpinned by causal laws, and hence they, too, are 'global', and depend on the character of a world as a whole. Causal connectedness, in contrast, is 'local'. The causal relation between two events does not depend on the overall pattern of events in the world around.

Causal laws and necessary and sufficient conditions, then, are 'global' whereas causation is 'local'. That is one reason why we resist analyses of

causation in terms of necessity or sufficiency of conditions. And the same reason extends to probabilistic analyses of causation. Whether an event boosts or depresses the probability of another event will be a global matter, not a local one. But a causal process counts as 'causal' entirely because of intrinsic, local, and not global, features.

Here is a more or less macroscopic example, modelled on a variety of microscopic, quantum-mechanical illustrations in the literature.[8] Consider an act of sexual congress between a male with low fertility, and a female. This act may cause conception and pregnancy to follow. Yet this act may not be a sufficient condition for these effects to follow. It may be largely a matter of luck that conception occurs. Nor need it be a necessary condition for pregnancy, that intercourse should occur with that male. Other males might be willing to take his place. In fact, by coupling with that male, the female may have depressed, rather than having raised, the chances of conception. If the female had not coupled with that male, we may suppose that some much more fertile male would swiftly have taken his place, in which case conception would have been almost certain.

In such a case, indeterminacy of outcome prevents the cause from counting as a 'sufficient' condition even in the sense of raising the probability of the effect; and the presence of another eager male prevents the cause from counting as a 'necessary' condition even in the sense that the effect would have been less likely in the absence of the cause.

It could be objected that the effect which results from the infertile male, which we have just called 'pregnancy', was in fact a specific pregnancy. Without the infertile male, the probability of *that* pregnancy would have been zero. If a different male had been involved, then a *different* effect would have resulted. Yet this objection is misplaced. It is a contingent fact that a different pregnancy is brought about by different males. Consider parthenogenesis with frogs' eggs, initiated by pin-pricking by a technician in a laboratory. The infertile male could be like a technician with a hand-tremour, while the more fertile male could be like a practiced, steady-handed technician waiting in the wings. (The ovist

[8] See W. Salmon, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, 1984.

preformation theory of generation of the eighteenth century gave the male a role very like that of such a technician.)

It might be further objected that, even though the low fertile male does not boost the probability of pregnancy compared to that with the more fertile male, nevertheless he does boost the probability compared to that with no male at all. Yet this again is only a contingent matter. Indeed, in the case of some species there is a probability of spontaneous parthenogenesis. The action of the male (or technician) may block any possibility of spontaneous parthenogenesis, and bestow a probability of generation which is *lower* than it would have been in the absence of interference. The act which causes the eggs to grow does not boost the probability of growth above what it would have, no matter *what* alternative this act is compared with.

We anticipate that some will respond that such examples are really cases of indirect causation, and that upon inspection each case will be found to contain a sequence of causal links, each of which does fit the probabilistic theory of causation. Yet such an appeal to mediating steps of necessary or sufficient conditions (or of boostings of probability) is highly speculative. The burden of proof is squarely on the defenders of such theories to establish that there are always such mediating steps. It is not easy to be confident that mediating steps of the right sort do always exist. The pervasive indeterminacy in fundamental physics suggests that, on the contrary, the mediating steps involved in a process like conception may all be ones which raise the very same problems over and over again, ones in which the causes fail to be necessary or sufficient or even to boost the probability of their effects. Furthermore, even if there *are* mediating steps, this seems not to be a consequence of the nature of causation itself, or to be part of what we *mean* when we say the sexual act caused the pregnancy.

Nevertheless, a debate conducted entirely at the level of speculative counterexamples is likely to be theoretically unproductive. Defenders of the modal theory, Lewis for instance, have considered examples like the ones described above, and have found them inconclusive. Thus we do not rest our case entirely on counterexamples. Our case against the probabilistic theory is also based on a prior theoretical consideration, and

we see the counterexamples not just as puzzles for the probabilistic theory, requiring *ad hoc* fine tuning, but as symptoms of deeper theoretical concerns.

In part, we rest our case as we have said on a construal of causation as a relatively localized, intrinsic physical transaction between two events. In part also, we rest our case on the role of causation in a wider explanatory context. We take causation to be part of the basic furniture of nature, and as such it functions as an input into the explanation of modalities. It is widely agreed that the best accounts of modalities make appeal to the framework of possible worlds. There is less agreement on how possible worlds are to be construed. Most of the details on the nature of worlds are unimportant here. What is important is only the direction of explanation between causation and the nature of worlds. We support theories which use causation as part of an account of what there is in any given possible world. Thus causation enters into the explanation of modalities, and in particular, into the explanation of 'necessary and sufficient conditions', and also of probabilities. Hence modal or probabilistic theories, even if they could be adjusted until they became extensionally correct, would nevertheless proceed in the wrong direction from an explanatory point of view. Causation is an input for theories of modality and probability, not an output.