Philosophy Compass 2/3 (2007): 508–516, 10.1111/j.1747-9991.2007.00084.x

Three Concepts of Causation

Christopher Hitchcock*
California Institute of Technology

Abstract

I distinguish three different concepts of causation: The *scientific* concept, or causal structure, is the subject of recent work in causal modeling. The *folk attributive* concept has been studied by philosophers of law and social psychologists. The *metaphysical* concept is the one that metaphysicians have attempted to analyze. I explore the relationships between these three concepts, and suggest that the metaphysical concept is an untenable and dispensable mixture of the other two.

In this review, I will discuss three distinct concepts of cause, which I will call the *scientific*, *folk attributive*, and *metaphysical* concepts. Corresponding to these three concepts, are three different putative kinds of causal relationship, scientific causation (which I will also call *causal structure*), folk attributive causation, and metaphysical causation. The scientific concept of causation is the one that lies at the heart of much recent literature on causal modeling. The folk attributive concept is the one used by ordinary folk in their attributions of praise and blame. This concept has been studied by social psychologists. Finally, the metaphysical concept of causation is the target of analysis in much of the philosophical literature on causation. I will discuss the relationships between the three different kinds of causal relation, and raise some skeptical worries about the metaphysical concept of cause.

The Scientific Concept of Causation

The scientific concept of causation is the focus of the recent literature on causal modeling (e.g. Spirtes, Glymour, and Scheines; Pearl). Physical systems possess a causal structure that can be represented by a *causal model*. Here is an illustration (based loosely on one given by Pearl ch. 3). Suppose that in a certain agricultural region, oat crops are threatened by eelworms. For this reason, farmers sometimes use fumigants to protect their crops. However, the fumigant may have a direct effect on the oat crop, and the eelworm population is independently controlled by a population of birds that prey upon the eelworms. To represent this system, we introduce a number of variables:

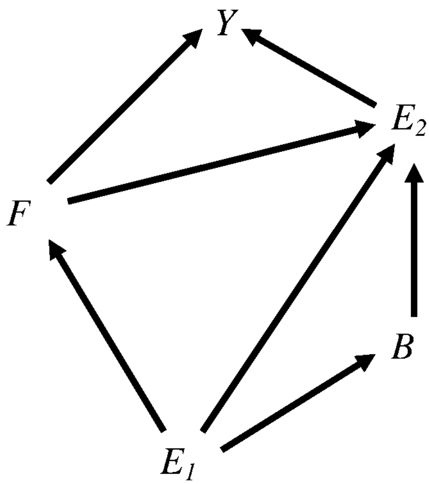

Fig. 1.

E_1 – the population of eelworms before the time at which the fumigant is (or would be) applied
F – the quantity of fumigant used (possibly zero)
B – the population of birds that prey on eelworms
E_2 – the population of eelworms after the time at which the fumigant was (or would have been) applied
Y – the yield of the oat crop

We might model this system using a series of linear equations:

CM1

$E_1 = U_{E1}$

$F = mE_1 + U_F$

$B = nE_1 + U_B$

$E_2 = pE_1 + qF + rB + U_{E2}$

$Y = sF + tE_2 + U_Y$

The Us represent 'error terms', the effects of variables that are not included in the model. In addition to the equations, we have a probability distribution *Pr* over the error terms. The qualitative structure of this model (with error terms omitted) is represented graphically in Figure 1.

The equations in the model represent *causal* structure and not merely correlations, because the equations provide information about the effects of hypothetical interventions, and about counterfactuals (specifically, non-backtracking counterfactuals in the sense of Lewis, 'Counterfactual Dependence'). For example, suppose that the initial eelworm population is $E_1 = e_1$. Then our model tells us that if we were to use quantity f of fumigant, the later expected eelworm population can be computed by replacing the equations for E_1 and F with $E_1 = e_1$ and $F = f$, respectively, and then solving the system of equations.[1] If, after the fact, we wonder what would have happened if we had used quantity f' of fumigant instead of f, our model tells us that the later eelworm population would have been different by a quantity $q(f' - f)$. Note, however, that not all of the equations that follow from our model can be used to answer questions about interventions or counterfactuals. For example, our model entails that B and F are correlated, since they share a common cause, E_1.[2] However, it would be wrong to use this relationship to infer that if we were to use quantity f' of fumigant instead of quantity f, the bird population would be different by $n/m(f' - f)$; in supposing an intervention to set the level of fumigant at f', we are supposing away the normal connection between fumigant level and initial eelworm population, so that the common cause structure is no longer intact. The equations that characterize the causal structure of the system, in contrast to those that merely describe correlations that exist in the system, are invariant under hypothetical interventions and counterfactual suppositions. (See Woodward for a detailed development of this theme.)

This conception of causation is scientific in the sense that the causal model generates a set of predictions about the system in question. In particular, the model entails a number of regularities; for example, as noted above, it predicts that in the absence of interventions, the bird population and quantity of fumigant used will be correlated according to a specific equation. The model also predicts the results of certain interventions on the system in question, such as the expected oat yield when the level of fumigant is set to f by an intervention. These predictions can then be compared with the statistics from the actual system, and the model can be confirmed or disconfirmed accordingly.

We make a number of observations about this scientific conception of causation. The first is that the causal structure is an objective feature of the system under investigation. In particular, the causal structure does not depend upon our interests or values. Of course, we are interested in this particular system because we are interested in oat yields; and we have a generally positive attitude toward high yields and a negative attitude toward eelworms. These interests and values may influence which systems we choose to investigate, but they do not affect the causal structures of those systems.

Second, it has become common in the philosophical literature to distinguish between singular or token causation, on the one hand, and general or type causation, on the other (see, e.g. Eells 'Introduction'). It might seem natural to reason that science deals with generalizations, and infer that the scientific conception of causation corresponds to general or type causation. This would be a mistake. For example, nothing prevents us from using CM1 to represent the causal structure of some particular oat field during some particular growing season. Indeed, it is easy to use causal models to represent causal structures in the kinds of scenarios typically discussed in the literature on singular or token causation. Consider a simple example of preemption: Assassin pours poison in King's coffee; if he hadn't done this, Backup would have; King drinks the poisoned coffee and dies. We can represent the causal structure of this scenario as follows:

A = 1 if Assassin poisons King's coffee, 0 if not

B = 1 if Backup poisons King's coffee, 0 if not

D = 1 if King dies, 0 if not

CM2:

$A = 1$

$B = \sim A$

$D = A \vee B$

This causal model captures all of the relevant counterfactuals that hold in the scenario. Of course, it will frequently be the case that we can't learn the causal structure of a complex system without making use of inductive evidence gathered from similar systems. But this epistemological truism doesn't entail that there is anything inherently general about the kind of causal structure that is represented by causal models such as CM1 and CM2.

The third observation is that the causal relations that comprise the causal structure of a system do not come neatly packaged in the form '*C* causes *E*', where *C* and *E* are events, facts, states of affairs, or properties. A causal model presents various relationships, and asserts that these are causal, in contrast to describing mere correlations. The relationships are causal because they support interventions and counterfactuals. One could perhaps paraphrase, e.g. the second equation of CM1 by saying 'eelworm populations cause fumigant levels', but such a paraphrase both strains the language and sacrifices quantitative information. When a causal model is used to characterize the causal structure of some specific episode, it will not in general be possible to 'read off' the relations of token causation from the causal model. For example, in the preemption scenario modeled by CM2, we naturally judge that Assassin's poisoning the coffee caused King to die, but we cannot read this causal relationship directly from the causal model CM2.

The Folk Attributive Concept of Causation

I take the term 'attributive' from Hart and Honoré, who distinguish the types of causal claim made in 'scientific' inquiries from those made in 'attributive' inquiries (24). It is the latter type of claim, they assert, that is of primary concern to 'the lawyer, the historian, and the plain man'. I have added the modifier 'folk' in order to emphasize that it is primarily the 'plain man's' conception of causation that I have in mind here. It is just this concept that the ‘plain man’ applies when he asserts, after the occurrence of some event, that some preceding event was the cause, or a cause, of it. Such causal attributions play a central role in assigning moral or legal responsibility for the event in question. What are the central features of the folk attributive concept of causation?

First, the causal relations in which this conception trades do have the form ‘*C* causes *E*’, where *C* and *E* are particular occurrences. Typically, *E* will be some event that has already occurred, triggering an ‘attributive inquiry’ into its causes. The causes will be earlier events judged to be responsible for the effect. Although we sometimes talk of people or objects as being causes, it is always understood that the person or object held responsible for the effect participated in some action or event that precipitated the effect.

Second, as Hart and Honoré note, causes in the folk attributive sense are distinguished from mere background conditions. Thus, for example, we count the careless tossing of the cigarette as a cause of the forest fire, whereas the oxygen in the atmosphere, while equally necessary for the occurrence of the fire, is relegated to the status of a mere background condition. A cause, on this conception, is some intervention or deviation from the standard course of events that results in some outcome distinct from what would have occurred in the absence of the deviation.[3]

Folk causal attributions have been studied in some detail by social psychologists. I wish to focus here on two features of the folk attributive concept of cause that have been revealed by these studies. The first is that ordinary subjects do not seem to view causation as an all-or-nothing affair: when invited to rate the extent to which they judge one event to be a cause of another on a numeric scale, they cheerfully choose intermediate values (see, e.g. the results reported in Alicke).

Second, folk causal attributions are influenced by normative factors, such as the existence of behavioral norms, or the intentions of agents. For example, in an experiment performed by Knobe and Fraser, subjects were presented a vignette in which two individuals, Lauren and Jane, both use the same computer. They can log on from separate terminals, but if they both log on at the same time, the computer will crash. In order to avoid this outcome, the company they work for establishes a policy allowing Lauren to use the computer in the morning, and Jane to use it in the afternoon. Then one morning, both Lauren and Jane log on to the computer, Jane doing so in violation of the company policy. The computer crashes. In this scenario, subjects are more strongly inclined to judge that Jane caused the computer to crash than that Lauren did. These results suggest that Jane’s violation of the company’s policy inclined subjects to judge that her action was a cause of the crash. Alicke presented subjects with a scenario in which an agent, John, was speeding on his way home. John is then involved in a traffic accident with another car. In one version of the story, John is speeding in order to get home before his parents so that he can hide an anniversary present that he left sitting out; in the other version, he is speeding in order hide a vial of cocaine that he left sitting out. Subjects were then asked to what extent John's speeding was a cause of the accident. Subjects judged that John caused the accident to a greater extent when he was speeding home to hide the cocaine. Thus John's intentions seemed to affect subjects' judgments of his causal role in the accident. These results, if taken at face value, suggest that causal attribution is not conceptually prior to the evaluation of moral responsibility, but is in fact more tightly bound up with judgments of moral responsibility: we are more strongly inclined to judge that an agent's actions caused some negative outcome when we judge the agent to be blameworthy in other respects. In this regard, ordinary causal attributions are not purely objective, since they depend, in part, upon our value judgments.

The Metaphysical Concept of Causation

The third concept of causation I call 'metaphysical', since it is discussed primarily in the metaphysics literature (here I take Lewis ('Causation'; 'Causation as Influence') and Hall as paradigmatic texts). This concept has features in common with both the scientific concept and the folk attributive concept. Like the folk attributive concept, applications of the metaphysical concept are typically retrospective; some event has occurred, and we assert that some other event was its cause. Moreover, the metaphysical concept, like the folk attributive concept, concerns causal relations of the form '*C* causes *E*', where *C* and *E* are particular events (or perhaps facts or states of affairs). However, the metaphysical concept, like the scientific, views causal relations as objective, existing independently of our evaluative judgments. Moreover, unlike the folk attributive concept, the metaphysical concept is 'nondiscriminatory' (Lewis, 'Causation' 162) or 'egalitarian' (Hall 227–8); it avoids 'invidious discrimination' (Lewis,'Causation' 162; Hall 228). Thus, for example, the metaphysical concept does not grade levels of causation, and it makes no principled distinction between causes and background conditions.

What is the relationship between the metaphysical concept of causation and the other two? A guiding assumption in much of the literature is that it should be possible to define the metaphysical relation of causation in terms of the causal structure of the scientific conception. For example, it is relatively straightforward to define the relation of counterfactual dependence (Lewis, 'Causation') or influence (Lewis,'Causation as Influence') in terms of causal structure; Lewis defines the causal relation as the ancestral of counterfactual dependence ('Causation') or influence ('Causation as Influence'). For other attempts to define the metaphysical concept of causation in terms of causal structure, see Hitchcock; Halpern and Pearl. This guiding assumption is connected with the presumed objectivity of metaphysical causation. Since causal structure is itself objective, a definition of metaphysical causation in

Philosophy Compass 2/3 (2007): 508–516, 10.1111/j.1747-9991.2007.00084.x

terms of causal structure would secure the objectivity of metaphysical causation.

How does the metaphysical concept of causation relate to the folk attributive concept? A predominant view (see, e.g. Lewis, 'Causation as Influence'; Hall) is that the metaphysical concept of causation is primary, while the folk attributive concept of causation is really just an artifact of the way in which we apply the metaphysical concept. For example, in talking about the causes of a fire, our interests may lead us to focus on a small number of causes and call them the primary causes of the fire. According to this view, background conditions, such as the presence of oxygen, are genuine causes, but such conditions are typically taken for granted, and are too numerous to mention explicitly. Thus pragmatic factors lead us to ignore or downgrade various causes, while emphasizing others. Similarly, if we are concerned with assigning blame for some bad outcome, it may be pragmatically inappropriate to point to the actions of someone who acted in accordance with accepted norms or out of benevolent motives, even if those actions were genuine causes of the outcome. We err when we mistake these pragmatic pressures for truth conditions, and judge that the concept of causation does not apply, or applies to a lesser degree, to certain events.

I want to suggest an alternative picture that opposes this standard view about the relationship between the metaphysical and folk attributive conceptions of causation: it may in fact be the metaphysical concept of causation that results from a mistake. Here is how the mistake arises. The folk make causal attributions when they assign praise and blame for various salutary and untoward outcomes. As philosophers, we naturally seek an objective basis for these folk causal attributions. In fact, the scientific concept of causation captures all there is to this objective core; beyond that, there are only our value-infected attributions of causal responsibility. As we have seen, the causal structure described by causal models does not yield causal relations with the grammatical form '*C* causes *E*', where *C* and *E* report, events, facts, or states of affairs. The objective causal structure of the world has a more complex grammar. However, in searching for the objective basis of our folk causal attributions, philosophers have mistakenly assumed that this objective basis will have the same logical structure as those attributions, that of a binary relation between events. This mistake has led philosophers to seek the objective basis of folk causal attributions in a metaphysical concept of causation, rather than in the causal structure of the scientific conception. This picture, if correct, would help to explain the failure of philosophers to capture the metaphysical concept of causation in objective terms. Metaphysical causation, it turns out, is an unstable compromise between the scientific and folk attributive concepts of causation: it seeks to retain the logical structure of the folk attributive concept while retaining the objectivity of the scientific concept.

This picture, which eliminates the metaphysical concept of causation, preserves all of the uses to which causal knowledge is put. We make use of causal knowledge to make predictions, and to guide our interventions in the world. For this purpose, it suffices to have knowledge of the causal structure of the scientific conception. We also make use of causal knowledge when we assign praise and blame for outcomes. For this purpose, the folk attributive concept of causation, value-laden as it is, is entirely appropriate. Thus, in characteristic applications of causal notions, we can make do without the metaphysical concept altogether.[4]

Short Biography

Christopher Hitchcock specializes in the philosophy of science, with a particular interest in the concept of causation. His articles have explored topics such as the role of contrast in causation and explanation, the analysis of singular causation using structural equation models, applications of causal concepts in the biological and social sciences, and the bearing of recent psychology on our understanding of causation. He is currently studying the relationship between causal and moral reasoning. His publications have appeared in journals such as the *Journal of Philosophy*, *Philosophical Review*, *Noûs*, *Philosophy of Science*, and *British Journal for Philosophy of Science*. He holds a B.A. in philosophy from Princeton University, as well as an M.A. in Mathematics and a Ph.D. in Philosophy, both from the University of Pittsburgh. He is presently Professor of Philosophy at the California Institute of Technology, and has also taught at Rice University.

Notes

* Correspondence address: Division of Humanities and Social Sciences 101–40, California Institute of Technology, Pasadena, CA 91125, USA. Email: cricky@caltech.edu.

[1] This gives us $E_2 = pe_1 + qf + rB + u_{E2} = pe_1 + qf + r(ne_1 + u_B) + u_{E2}$.
[2] More precisely, they are related according to the equation $B = n/m(F - U_F) + U_B$.
[3] Barring the presence of backup causes which would have produced the same effect, had the actual cause been absent.
[4] Thanks to Josh Knobe and Jiji Zhang for helpful comments on earlier drafts of this article.

Works Cited

Alicke, M. 'Culpable Causation'. *Journal of Personality and Social Psychology* 63 (1992): 368–78.
Driver, J. 'Attributions of Causation and Moral Responsibility'. *Moral Psychology, Vol. 2, The Cognitive Science of Morality*. Ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press, forthcoming.
Eells, E. *Probabilistic Causality*. Cambridge: Cambridge UP, 1991.
Hall, N. 'Two Concepts of Causation'. *Causation and Counterfactuals*. Eds., J. Collins, N. Hall, and L. A. Paul. Cambridge, MA: MIT Press, 2004. 225–76.
Halpern, J. and J. Pearl. 'Causes and Explanations: A Structural-model Approach – Part I: Causes'. *British Journal for the Philosophy of Science* 56 (2005): 843–87.
Hart, H. L. A. and A. Honoré. *Causation in the Law*. 2nd ed. Oxford: Oxford UP, 1985.
Hitchcock, C. 'The Intransitivity of Causation Revealed in Equations and Graphs'. *Journal of Philosophy* 98 (2001): 273–99.

Knobe, J. and B. Fraser. 'Causal Judgment and Moral Judgment: Two Experiments'. *Moral Psychology, Vol. 2, The Cognitive Science of Morality*. Ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press, forthcoming.

Lewis, D. 'Causation'. *Journal of Philosophy* 70 (1973/1986): 556–67. Reprinted in D. Lewis. *Philosophical Papers, Vol. 2*. Oxford: Oxford UP, 1986. 159–72. [Page references are to the 1986 volume.]

——. 'Causation as Influence'. *Journal of Philosophy* 97 (2000): 182–97.

——. 'Counterfactual Dependence and Time's Arrow'. *Noûs* 13 (1979/1986): 455–76. Reprinted in D. Lewis. *Philosophical Papers, Vol. 2*. Oxford: Oxford UP, 1986. 32–52.

——. *Philosophical Papers, Vol. 2*. Oxford: Oxford UP, 1986.

Pearl, J. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge UP, 2000.

Sinnott-Armstrong, W., ed. *Moral Psychology, Vol. 2, The Cognitive Science of Morality*. Cambridge, MA: MIT Press, forthcoming.

Spirtes, P., C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press, 2000.

Woodward, J. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford UP, 2003.