HAL
archives-ouvertes.fr

The Landscape of Causation

Max Kistler

▶ **To cite this version:**

Max Kistler. The Landscape of Causation: Review-essay on L.A. Paul and Ned Hall, Causation : A User's Guide, Oxford Université Press, 2013. Metascience, Springer Verlag, 2014, 23 (3), pp.497-504. <10.1007/s11016-014-9906-7>. <hal-01292012>

HAL Id: hal-01292012
https://hal.archives-ouvertes.fr/hal-01292012

Submitted on 22 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Max Kistler
Université Paris 1 Panthéon-Sorbonne, CNRS, ENS
UMR 8590 IHPST - Institut d'Histoire et de Philosophie des Sciences et des Techniques
mkistler@univ-paris1.fr

The Landscape of Causation. Review-essay on L.A. Paul and Ned Hall, *Causation : A User's Guide*, Oxford Université Press, 2013, *Metascience*, published online 19 June 2014. DOI: 10.1007/s11016-014-9906-7.

L. A. Paul and Ned Hall's book makes an original and important contribution to the philosophical debate on causation. Their aim is not to construct a theory of causation but "to sketch a map" of the "landscape" (p. 1) constituted by a rich set of problem cases and various theories of causation devised to account for them.

Chapter 1 presents the scope and aim of the book, justifies the method of evaluating theories of causation by exploring whether they are refuted by counterexamples, and provides an overview of the rest of the book.

Chapter 2 justifies the choice of problems and solutions considered in the book. Several assumptions are made from the beginning. Causation is taken to be a relation among particular events rather than types of events. A "reductionist outlook" is adopted, according to which facts about what causes what are determined by i) facts about what happens and ii) fundamental laws. These assumptions provide the background for a "minimal sufficiency account" of causation (p. 16), which is a version of a regularity account. The other approaches that are introduced and examined in light of their capacity to cope with problem cases include the following: probabilistic accounts, transference accounts, and counterfactual accounts. Several versions of the counterfactual account get special attention: 1) Lewis' (1973) theory of chains of non-backtracking counterfactual dependence, 2) Lewis' (2000) account of causation as influence, 3) accounts of "de facto dependence", versions of which have been put forward by Yablo (2004) and Hitchcock (2001), and finally 4) contrastivism, according to which causal relations have the form "C rather than C* causes E rather than E*". But Paul and Hall provide more than a useful overview of the literature. They analyze the aims that the various approaches fix themselves or, in the authors' own words, the "conflicting motivations and conflicting presuppositions about the very point of providing a philosophical account of causation" (p. 1).

Their own conception of giving an account of causation is expressed in a list of rules. An account of causation is satisfactory if 1) it is reductive and avoids circular use of causal notions in the analysis, 2) it is metaphysically conservative and avoids the postulation of "extravagant" entities, such as negative events to account for causation by omission, 3) it is conceptually parsimonious and tries to analyze causation as directly as possible in terms of fundamental categories, 4) it applies not only to our actual world but also to other possible worlds, 5) it leaves open the possibility that some causal intuitions are wrong.

Chapters 3 to 5 evaluate in rigorous detail how several approaches, the most promising ones according to Paul and Hall, cope with problem cases. Chapter 3 explores various types of "redundant causation", such as preemption and overdetermination, which have been the focus of an important part of the literature on counterfactual approaches to causation. Chapters 4 and 5 explore causation by omission and two types of situation in which it is unclear whether causation is transitive, i.e. double prevention and switching. The last chapter, Chapter 6 contains concluding remarks.

The book contains helpful indices, both for subjects and names, and a useful compilation of all 49 so-called "neuron diagrams" that are used throughout the book to represent test cases. There are a few typos, some of which cause the reader trouble, such as when "agreement" (p. 11) replaces "disagreement", or when fig. 13 is mentioned (p. 92) as an example for something that is not true of it

In chapter 2.3., Paul and Hall make a subtle and original contribution to understanding the existence of several paradigms of research on causation. They highlight differences in the aims that research programs on causation fix themselves and in the criteria they consider appropriate for judging their success. Woodward (2005), for example, analyzes relations between causation and intervention, and among a number of different causal concepts. However, his theory is non-reductive because it analyzes the notion of intervention in causal terms. Paul and Hall distinguish between stipulative definition, conceptual analysis, and ontological reduction, which corresponds to their own conception of the task. However, the panorama is incomplete since two important options are left out.

According to the first omitted option, the task of giving an analysis of causation consists in analyzing the methodology underlying the scientific research for causes. Such a conception underlies some of the most important philosophical approaches to causation: Woodward's interventionist analysis takes up Pearl's (2000) and Spirtes, Glymour and Scheines' (2000) work on causal modeling in terms of structural equations and acyclic graphs. The aim of this research is to make explicit the algorithms used by scientists who look for causal dependence among variables in models of complex systems, like those used in econometrics or epidemiology. Similarly, the paradigm of analyzing causation in terms of conditional probability can be seen as making explicit, clarifying and improving, the methodology of the scientific research for causes in the social sciences. (The deductive-nomological model was taken to make explicit the concept of cause underlying causal explanation in the physical sciences.)

The reason Paul and Hall give for neglecting probabilistic accounts is that "our focus is on causation in the deterministic domain" (p. 23). It is not clear what they mean by the "deterministic domain". Maybe it is meant to exclude fundamental physical indeterministic processes, and to include such commonsense events as explosions. In that case, their reason for neglecting accounts of causation in terms of conditional probability is weak because such accounts have been devised for such domains as economy and sociology, and should as well apply to such events as explosions. The probabilistic approach is applicable to such events as well as to economic and sociological events, not because they are somehow intrinsically indeterministic, but because the situation in which they occur is incompletely known and described.

A second neglected suggestion is Menzies' (1996) suggestion to apply functional analysis to causation. Causation is a natural kind of relation that fills, in the actual world, a certain conceptual role characterized by a set of descriptions: causation can serve as a means to an end, raises probability, is asymmetrical etc. Being partly conceptual and partly empirical, such a conception does not fall into any of the categories distinguished by Paul and Hall. The discovery of the nature of causation is guided by conceptual constraints corresponding to the role causation plays, but it cannot be completed without the contribution of a posteriori scientific information. Special relativity provides such a constraint on what fills, in the actual world as it is explored by science, the role of causation: According to special relativity, there are no causal processes exceeding the speed of light.

Paul and Hall neglect science altogether, which is surprising because they announce, at the beginning of the book, that they see their task as that of "developing scientifically informed notions of 'cause'" (p. 2). They do not fulfill this promise because taking account of science is incompatible with their "rule four: 'thou shalt not be an ontological wimp'" and "appeal, in one's reduction, to facts too specific to one's own world" (p. 40). An account cannot both be informed by science, which enquires only into the actual world, and apply to possible worlds where our actual laws of nature do not obtain, such as fictional worlds in which magicians can provoke effects instantaneously at arbitrary distance, or across arbitrary time intervals. Rule four shows that Paul and Hall take the task to be the analysis of our concept of causation – that is at least partly independent from both methods and results of science – rather than of causation itself, as it is in the actual world.

Throughout the book, the main contenders for an acceptable account are those based on counterfactual dependence and minimal sufficiency. The structure of the book follows from this choice: The bulk of the volume carefully examines the potential counterexamples to such approaches, in terms of "redundant causation, such as various types of preemption and overdetermination (Chapter 3), omissions (Chapter 4), and cases that threaten transitivity (Chapter 5).

The "minimal sufficiency account" is a version of the idea underlying regularity accounts, according to which causes are, in the circumstances, lawfully sufficient for their effects. According to this account, "C is a cause of E iff C belongs to a set of contemporaneous events that is minimally sufficient for E" (p. 16). However, it faces a fundamental difficulty. Without a circular restriction to "causal laws", nomic sufficiency is not a sufficient condition for causation. The pendulum's period is nomically sufficient for its length, but not a cause of its length.

Although Paul and Hall give "special attention to counterfactual and related analysis of causation" (p. 1-2), they distance themselves from David Lewis' conception of the task as pure conceptual analysis, which aims at making explicit the conceptual structure implicit in commonsense intuitions and causal statements expressed in ordinary language. For such an approach, the main criterion for evaluating a philosophical theory of causation is the agreement of the result of applying the theory to various situations with intuitions of commonsense. Paul and Hall think that the analysis of causation must start with the analysis of the concept implicit in commonsense intuitions, but they take their task to belong to revisionary and not descriptive metaphysics, so that commonsense should not necessarily have the last word: "running afoul of common sense is not an automatic disqualifier" (p. 3). However, what else if not science might justify an account that contradicts intuition?

Throughout the whole book, Paul and Hall apply a methodology appropriate for conceptual analysis: they examine accounts that do not require any scientific ingredients, and evaluate them by the "method by counterexample" (p. 249), which consists in checking whether their verdicts on a number of situation schemata agree with the verdict of common sense intuition. The only reasons they accept that a theory may contradict commonsense intuition is that there are types of situation - especially involving pre-emption, overdetermination, and omission - in which common sense either has no clear intuitions or has different mutually conflicting intuitions on the very same situation.

Paul and Hall claim that cases of redundant causation only threaten the claim that counterfactual dependence is *necessary* for causation: In the presence of a backup cause, the effect does not counterfactually depend on its cause. However, "as a sufficient

condition on causation" (p. 16), they take counterfactual dependence to be "an excellent starting point for a full blown analysis of causation" (p. 16). This optimism is surprising given that the authors acknowledge that there is no satisfying answer to the objection that there are true backtracking counterfactuals, according to both intuition and scientific considerations. Given that causes sometimes depend counterfactually on their effects, counterfactual dependence cannot be sufficient for causation. Paul and Hall judge that Lewis' defense of the counterfactual dependence account against this objection, in terms of miracles, "is wholly unsatisfactory" (p. 46). They suggest an alternative method of evaluating counterfactuals, which they adapt from Maudlin (2007), and claim that it provides a non-circular justification of the thesis that there are no true backtracking counterfactuals. Here is the "Maudlin recipe" for evaluating the counterfactual "if C had not occurred, E would not have occurred". Construct a description that differs from the description of the actual world at time t, when C occurs, only by omitting C itself. Then "evolve the resulting state forward in time, in accordance with the actual laws of nature" (p. 48). If this leads to E, the counterfactual is false, if not, it is true. Paul and Hall take this recipe to justify the claim that backtracking counterfactuals are always false. However, it seems to deliver the opposite result: Fundamental physical laws are time reversible, so that, as the authors themselves explain (p. 61), the recipe can also be applied to evaluate backtracking counterfactuals. It is false that this "recipe has the 'non-backtracking' feature that…it needs to have in order to have a chance of undergirding a successful account of causation" (p. 49). However, although no good reason has been given for judging that backtracking counterfactuals are always false, counterfactual approaches are explored in the rest of book presupposing such a "non-backtracking reading of the counterfactual" (p. 71).

Chapter 3 explores several sophisticated theories of counterfactual dependence and their capacity to deal with difficult cases of preemption: Hitchcock's (2001) and Yablo's (2004) accounts of so-called "de facto dependence", and the authors' own "intrinsicness thesis" (p. 127), which takes up Lewis' concept of "quasi-dependence". The idea is to justify that C causes E in terms of counterfactual dependence in an indirect way, when, in the presence of a backup, E does not counterfactually depend on C. The idea behind both quasi-dependence and intrinsicness is that C nevertheless causes E if there exists another structure (containing counterparts C' and E' of C and E and of all intermediate causes between C and E), sharing its intrinsic character with the original structure and in which E' depends counterfactually on C'. True, it seems to be a "deeply intuitive idea" that it depends only on the intrinsic nature of the structure of events between C and E whether C causes E, and not on their extrinsic relations to events that are not involved in the causation of E. The problem is that without a definition of "intrinsicness" (the authors admit having none, p. 124), it is difficult to avoid the suspicion that the condition of intrinsicality has no non-circular content. 1) It is only when the structure of the causes of E is already known that one can identify the structure of events whose intrinsic nature determines whether C causes E. 2) Intrinsicness may seem deeply intuitive only because it is easily confused with locality. It is intuitive that causes act only locally. However, if at some world C causes E at a temporal and/or spatial distance, the intrinsicness thesis says that the causal influence of C on E is independent from what happens at the local surroundings of C and E, which is not intuitive at all.

The treatment of transference theories is surprising. Several reasons are given to reject them. However, the authors keep considering them throughout the book, judging

them even sometimes superior to its rivals (p. 212). If they are so clearly refuted, why keep taking them seriously as contenders?

The first argument against transference theories is that they "suffer from a surprising lack of ambition" (p. 55). Transference theories are indeed intended to apply only to causation as it is in the actual world, as opposed to other possible or merely conceivable worlds. According to transference theories, changes in the occupation of the cells that make up Conway's "Game of Life" are not causal. The cells correspond to locations in an abstract space and the changes in their occupation are determined by a mathematical rule. Maybe there are possible worlds that have such structures. That the changes in the occupation of cells are causal processes rather than mathematically determined sequences of formal structures is no neutral judgment that can be used to judge the adequacy of transference theories. They may be intuitively causal, but the authors' judgment based on such an intuition ("there seem to be causal relations in such a world" p. 56) can only be used to evaluate theories that aim at analyzing common-sense intuition, not theories that aim at finding out what causation is in the actual world. The former enterprise may be more ambitious in one sense because it applies not only to the actual world but also to those non-actual possible worlds about which we have causal intuitions. In another sense, it is less ambitious, because it does not aim at what causation is (in the actual world), but only at what common sense intuition takes it to be.

Paul and Hall take transference theories to be inadequate for at least two other reasons. 1) Transference of an amount of energy or of some other conserved quantity does not seem to be sufficient for causation because a process or event C can transfer something to an event E without intuitively being among its causes. To take an example that is used throughout the book: Suzy throws a rock that hits and smashes a bottle. Billy also throws a rock at the same bottle but is a tiny bit slower so that his rock gets to the place where the bottle was only after it has been smashed by Suzy's rock. There may well be air molecules that carry momentum or energy from Billy's rock to the bottle although Billy's rock is not among the causes of the breaking of the bottle. The thesis that transference is sufficient for causation can be defended against this objection by distinguishing between causation as such and a more restrictive notion of causal "proportionality" (Yablo 1992, Woodward 2010) or "causal responsibility". The air molecules make Billy's rock causally influence the breaking of the bottle. However, this is no good reason to call Billy's rock a cause of the bottle's breaking insofar as one uses the word "cause" to mean "proportional cause": The air molecules are not proportional to the effect as a breaking of a bottle, in the sense that, without them, the bottle would still have broken, though not in the same way. The molecules are proportional to a much more precise type of breaking, although one in which no one is normally interested.

(Distinguishing causation simpliciter from proportional causation suggests a promising way of solving the problem of the overdetermination of behavior by mental and physical causes, discussed p. 155-61. However, Paul and Hall do not mention authors, such as Yablo (1992) and List and Menzies (2009), who have used such distinctions in the context of solving this problem. The same distinction would also provide a simple means of analyzing the difference between the complex "black-box"-cases analyzed on pp. 162-7.)

2) Cases of omission and prevention seem to show that transference is not intuitively necessary for causation. To take an example from the book (p. 174), David's desire for coffee causes him to take a sip. Steffi stands nearby and does nothing. However, had she gestured wildly, she would have knocked over David's cup, so that he wouldn't have drunk. One may have the intuition that Steffi's omission of gesturing is

one of the causes of David's drinking from his cup. However, the necessity of transference for causation may be defended by pointing out that the intuition is open to different interpretations. True, if Steffi had not stood still but had instead wildly gestured, David would not have drunk. However, it is not obvious that this is the intuition that her standing still is *causing* David's sip. As far as intuition is concerned, it may be a case of non-causal dependence. The fact that David's sip depends on Suzy's standing still is only equivalent to the fact that Suzy's standing still causes David's sip if it is presupposed that counterfactual dependence is equivalent to causation. The intuition provides no non-circular justification of the thesis that omissions are causes, and therefore of the thesis that transference is not necessary for causation.

In Chapter 4, Paul and Hall analyze the inextricable difficulties all accounts face if omissions are taken to be cases of causation, as indeed they intuitively appear to be. So why not take seriously arguments for taking omissions to be cases of non-causal dependence rather than causation (a distinction common-sense intuition does not make)? One such argument, mentioned on p. 183/4, is that the intuitive judgment of whether an omission is a cause is determined by normative considerations: The gardener's failure to water the flowers is a cause of their death, but Putin's failure to water them is not. Unless norms are given a metaphysical status, this shows that the difference between the relation of the gardener's failure to the plants' death (dependence) and the relation of Putin's failure to the plants' death (non-dependence) is not metaphysical, and thus that it is not a difference in causation. Sometimes Paul and Hall themselves suggest that omissions cannot be causes, e.g. when they say that "the wall doesn't do anything, so it can't be a cause" (p. 189).

The aporetic end of Chapter 4 is characteristic of Paul and Hall's work: They admit being "unable to give a decently unified or complete account of" (p. 214) omission "using any of the standard approaches to causation", but refrain from drawing the conclusion that omission is not causation. The only reason offered for being "unable to dismiss causation involving omissions" (p. 214) is that intuition assimilates omission to causation. However, it might be wiser in this case to follow their own "rule five: 'Thou shalt not enshrine intuitions.'" (p. 41)

Paul and Hall's book is a rich and sophisticated piece of philosophical work. No doubt it's detailed analyses make a precious contribution to the fascinating philosophical debate on causation. However, the book's subtitle "a user's guide" adequately describes the book only in part. I would not recommend Chapters 3-5 to readers not already familiar with the philosophical literature. Some passages are terribly difficult. In one of the neuron diagrams analyzed on pp. 135-9, one neuron can fire with different intensities, one of which is "I", the inhibiting intensity; and it can fire in different shades of grey, one of which is the "triggering shade". Although they provide a table describing what happens in several cases, the situation is hard to understand. The crucial case, their "test case", in which the neuron fires *both* with the triggering shade and the inhibiting intensity, does not figure in the table. They stipulate that in this case, and if a second neuron D fires, then E fires. All right, but how can they say on the next page that the label "'inhibiting intensity' for firing with intensity I seem(s) entirely appropriate" (p. 137), given that, in the test case, C fires with intensity I and nevertheless does *not* inhibit E from firing?

The analysis of Yablo's account of de facto dependence is bewildering. Indeed, the crucial distinction it uses between "good" and "bad" (p. 83) choices of the facts that are held fixed when judging counterfactual dependence among variables is so unintuitive and so difficult to apply that it might well be a case of what Paul and Hall call

"unproductive epicycling" (p. 251), given that the conclusion of its careful evaluation is that "it is not obvious that (Yablo's account) can be developed in a way that produces a precise, successful account" (p. 89)[1].

References

- Hitchcock, C. (2001), The Intransitivity of Causation Revealed in Equations and Graphs, *Journal of Philosophy* 98, p. 273-99.
- Lewis, D. (1973), Causation, *Journal of Philosophy* 70, p. 556-67, repr. in D. Lewis, David, *Philosophical Papers, vol. II,* New York, Oxford University Press, 1986, p. 159-172.
- Lewis, D. (2000), Causation as Influence, *Journal of Philosophy* 97, p. 182-197.
- List C. and P. Menzies (2009), Non-Reductive Physicalism and the Limits of the Exclusion Principle, *Journal of Philosophy* 106, p. 475-502.
- Maudlin, T. (2007), A Modest Proposal Concerning Laws, Counterfactuals, and Explanation, in T. Maudlin, *The Metaphysics Within Physics*, Oxford: Oxford University Press, p. 5-49.
- Menzies, P. (1996), Probabilistic causation and the pre-emption problem, *Mind* 105 p. 85-117.
- Pearl, Judea (2000), *Causality. Models, Reasoning, and Inference*, Cambridge, Cambridge University Press.
- Spirtes, Peter, Glymour, Clark and Scheines, Richard (2000), *Causation, Prediction and Search*, Second edition, Cambridge (Mass.), MIT Press.
- Yablo, S. (1992), Mental Causation, *Philosophical Review* 101, p. 245-80.
- Yablo, S. (2004), Advertisement for a Sketch of an outline of a Proto-Theory of Causation, in J. Collins, N. Hall, and L.A. Paul, eds., *Causation and Counterfactuals*, Cmabridge, MA: MIT Press, p. 119-37.
- Woodward, J. (2005) *Making Things Happen*, Oxford University Press.
- Woodward, J. (2010), Causation in biology : stability, specificity, and the choice of levels of explanation, *Biology and Philosophy* 25, p. 287-318.

[1] I thank Andrew McFarland for improving my English.