

David Lewis' Theories of Causation and their Influence
(Forthcoming in *Cambridge History of Philosophy*, ed. Kelly Becker)
Sara Bernstein

David Lewis' metaphysics of causation set the stage for many contemporary approaches to the topic and laid the groundwork for debates on related dependent philosophical concepts, including interventionist theories of causation, causal modelling, grounding, and the role of laws in metaphysics. This article will give an overview of Lewis' work on causation, and trace the influence of Lewis' approach through recent developments in the metaphysics of causation. Gameplan: first I will give a summary of Lewis' views on causation, as well as various well-known challenges they faced. Next I will summarize the well-known attempts to respond to these problems that dominated the causation literature for many years after Lewis' early work. Then I will give an overview of the myriad and widespread topics that share an ancestor in Lewis' theories of causation. Finally, I will examine the influence of Lewis' approach on present-day "hot topics" in metaphysics.

Lewis on Causation

The linchpin of Lewis' views on causation is the *counterfactual*. Counterfactuals (roughly, statements of the form "If c hadn't occurred, e wouldn't have occurred") are intrinsically interesting for Lewis, but also unite several strands of his system: his possible worlds semantics (which postulates a similarity metric across possible worlds), his modal realism (which postulates infinite, real world-sized possibilities), and his views on causation, which will be the focus of our discussion. His unified system and his view of counterfactual specifically set the stage for the next half-century of metaphysics.

Before Lewis' counterfactual theory, postwar approaches to causation were generally concerned with formulating causation in terms of necessity and sufficiency of causes for their effects. Mackie (1965) proposed an analysis of causation in terms of INUS conditions, or Insufficient but Necessary parts of conditions which are Unnecessary but Sufficient for their outcomes. Hempel's (1965) covering-law model of explanation saw causes as subsumed under laws of nature, with the latter serving as

premises in deductively sound arguments about worldly *explananda*. Causes, on this approach, are inexplicably bound up with natural laws, which are necessary conditions for successful causal explanations. Resistance to these approaches centered around the required involvement of lawful regularity in successful causal claims: one can have a successful causal explanation, it seems, without involving broad natural laws. Scriven (1962) suggested that causal claims need not always be considered instances of lawful necessities.

Drawing on the idea of a logical connection between cause and effect that had to be something besides necessity, sufficiency, or lawful regularity, Lewis' "Causation" (1973) became the seminal statement of the counterfactual approach. In that piece, Lewis argues that causation is the ancestral of the relationship of counterfactual dependence: *c* is a cause of *e* if it is the case that for two actual events *c* and *e*, had *c* not occurred, *e* would not have occurred. For example: suppose that Billy throws a rock at a window, causing it to shatter. The counterfactual "If Billy had not thrown the rock through the window, the window would not have shattered" is true. (We are not to admit *backtracking* readings of counterfactuals, such as "If the window shattered, it's because the rock shattered it", owing to their falsity on ordinary interpretations, Lewis claims.) For Lewis, evaluating counterfactuals requires appealing to possible worlds, as distilled in the following rule:

"If A were the case, C would be the case" is true in the actual world if and only if (i) there are no possible A-worlds; or (ii) some A-world where C holds is closer to the actual world than is any A-world where C does not hold.¹

Closeness of worlds is to be judged in terms of *comparative overall similarity* to the actual world based on similarities between those worlds and the actual one. Additionally, so that causation is transitive, causation is the ancestral of counterfactual dependence: there is a string of counterfactual dependencies between *c* and *e*.

Lewis' 1973 view, however, faced obvious counterexamples in different kinds of *redundant causation*, or cases in which there are multiple sufficient causes to bring about

¹ This formulation of the rule is drawn from Menzies (2014).

an outcome. Suppose that along with Billy, Suzy threw her rock at the window at almost the same time. Then the counterfactual “If Billy had not thrown his rock, the window would not have shattered” is false, as well as the counterfactual “If Suzy had not thrown her rock, the window would not have shattered.” For either is false in virtue of the fact that the other person’s rock would have shattered the window. On the simple counterfactual account, the shattering of the window is left entirely without causes: an unacceptable result.

There are multiple sorts of redundant causation, and each impacts the simple counterfactual theory in a slightly different way. There is a case of *early preemption* when there are multiple would-be causes for an outcome, but one causes preempts the other would-be causal process from running to completion. To use a canonical example: Billy and Suzy are each poised to throw their rocks at a window. Suzy throws her rock at the window artfully. Billy, who was poised to throw his, becomes demoralized and puts his rock down.

In this sort of case, had one cause not brought about the effect, the other cause would have, even though the actual process of the would-be cause never ran to completion. Taking causation as the *ancestral* of counterfactual dependence, as Lewis does, solves the problem with early preemption, for there is a chain of counterfactual dependencies running from Suzy’s rock to the window’s shattering, but not from Billy’s rock to the window’s shattering. The simple counterfactual account can handle this and similar cases of early preemption.

It does not fare so well, however, in cases of late preemption and overdetermination. There is a case of *late preemption* where, roughly, there are multiple causes sufficient to bring about an outcome, and both causal processes run to completion. In many cases of late preemption, the preemption cause preempts the would-be cause by bringing about the effect first. For example, suppose that Billy and Suzy each throw rocks at the window. Suzy’s rock strikes the window and breaks it, and Billy’s rock flies through the space where the window used to be. Here, there are separate chains of counterfactual dependencies running from both Suzy’s rock and from Billy’s rock. Similarly with cases of *overdetermination*, in which there are multiple causes sufficient to bring about an outcome in the way it occurs. Suppose, for example, that Billy and Suzy

each throw their rocks at the window and both rocks hit the window at the same time. Either rock along could have shattered the window in roughly the way it occurred. Each causal process is a counterfactual dependence-breaking backup for the other, and on the simple counterfactual account, they are not to be causally distinguished.

Responses to the Counterexamples

The period after these well known counterexamples was concerned with emending the counterfactual account of causation in order to best shore up the theory. Yablo's (2002) attempt came in the form of the *de facto dependence* theory of causation, which holds that *e* de facto depends on *c* if had *c* not occurred, and had other factors been held fixed, then *e* would not have occurred. De facto dependence is a theory built to handle preemption cases, since holding fixed facts like whether or not Billy threw his rock in addition to Suzy yields the right result about counterfactual causation. The "holding certain things fixed" strategy laid the groundwork for the causal modelling approach, which sought to formalize such relationships. I discuss this approach further below.

Lewis' own (2000) attempt at shoring up the counterfactual theory proposed a promising set of bells and whistles on the original theory. As opposed to the "whether-whether" dependence of the simple counterfactual approach (that is, whether *e* occurs depends on whether *c* occurs), Lewis proposes "whether-when-how" dependence of *e* on *c*: whether, when, and how *e* occurs depends on whether, when, and how *c* occurs. More formally, causation is a matter of counterfactual covariation between modally fragile alterations on an actual event *e* and an actual event *c*. For example: suppose that Billy throws his rock in a particular way. Call that actual event *c*. Now suppose that the window shatters in a particular way because of the rock throw. Call that actual event *e*. A modally fragile alteration on *c* is just a slight variation of the way *c* occurs: for example, Billy throwing the rock slightly earlier. A modally fragile alteration on *e* is just a slight variation of the way *e* occurs, for example, the window shattering slightly earlier. To check for causation, we check to see if there is a pattern of counterfactual covariation between these modally fragile alterations. For example, is it the case that if Billy had

thrown his rock slightly differently, the window would have shattered slightly differently?

Lewis intended this more complicated account to handle cases of redundant causation in the following way: an event that influences an effect (in his technical sense of “influence”) is considered more of a cause of an outcome than events with less influence. Consider the bugaboo of the 1973 counterfactual account, a case of late preemption: Suzy throws her rock at the window, shattering it, and Billy’s rock, thrown only slightly earlier, flies through the space where the window used to be. Altering Suzy’s rock-throw—changing when and how it occurs—will change when and how the window shatters more than altering Billy’s rock. Suzy’s rock, the preempting cause, thus exhibits more influence than Billy’s rock, the preempted cause.

These results had several interesting outcroppings. First, Lewisian influence causation is arguably a matter of degree rather than “on or off”: one event can cause something to a greater degree than another. Second, even “trace” causes (for example, a dog wagging his tail near the shattering of the window) exhibit slight influence over outcomes, since altering how trace events occur slightly alters the way outcomes occur.

Finally, the influence account is arguably better able to handle cases of *trumping preemption*, cases in which there are multiple sufficient causes for an outcome, and one “trumps” the other according to a pre-existing rule or law with no discernable difference in causal process. In Schaffer’s famous (2000a) example:

“Imagine that it is a law of magic that the first spell cast on a given day match the enchantment at midnight. Suppose that at noon Merlin casts a spell (the first that day) to turn the prince into a frog, that at 6:00 PM Morgana casts a spell (the only other that day) to turn the prince into a frog, and that at midnight the prince becomes a frog. Clearly, Merlin’s spell (the first that day) is a cause of the prince’s becoming a frog and Morgana’s is not, because the laws say that the first spells are the consequential ones.”

According to Schaffer, this is not a case of late preemption because both causal processes “run to completion,” and it is not a case overdetermination because only one process, Merlin’s spell, is the cause of the enchantment.² Lewis holds that the influence

² There is some controversy over whether trumping should be viewed as a kind of preemption or a kind of overdetermination: see e.g. Bernstein (2015) and Hitchcock (2011).

account can handle Schaffer's canonical case of trumping preemption: the transmutation is more sensitive to variations in Merlin's spell than in Morgana's. However, it has been pointed out (Strevens 2003, Collins 2007) that similar cases of trumping preemption can be constructed which include effects that are insensitive to variations on the putative cause. It remains a matter of some controversy how well Lewis' later account handles such cases.

2000-Present: The Causal Modelling Era

In the early 2000s, attention began to shift from the Lewisian armchair methodology to approaches that sought to precisify counterfactual relationships within formal structural equations models. Central figures of the causal modelling movement include Pearl (2000), Hitchcock (2001), and later Schaffer (2016). Generally, causal models have the following key ingredients: *variables* which represent the occurrence, non-occurrence, or nature of token events, *values* assigned to those variables, and *structural equations* which represent the causal relations among those elements. A causal model is often represented in an ordered triple $\langle U, V, E \rangle$, where U represents values of the variables not in the causal model (*exogenous* variables), V represents the variables within the causal model (*endogenous* variables), and E represents one or more structural equations modelling the causal relationships.

Causal models are best understood by example. Consider the late preemption case in which Suzy's rock preempts Billy's rock. That scenario can be represented in the following model, modified from Menzies (2004a):

Suzy Throws (ST) = 1 if Suzy throws a rock, 0 if not.

Billy Throws (BT) = 1 if Billy throws a rock, 0 if not.

Suzy Hits (SH) = 1 if Suzy's rock hits the intact window, 0 if not.

Billy Hits (BH) = 1 if Billy's rock hits the intact window, 0 if not.

Bottle Shatters (BS) = 1 if the window shatters, 0 if not.

Here, the exogenous variables are ST and BT: they are not “within the control” of the causal modeler or within the causal model itself. The endogenous variables are determined by the causal modeler. There is a structural equation for each endogenous variable. The structural equation often models counterfactual dependence or independence. Drawing again on Menzies’ setup, the causal relationships in the above model can be represented in the following way:

$$\textit{Suzy Hits} = \textit{Suzy Throws}$$

$$\textit{Billy Hits} = \textit{Billy Throws} \ \& \ \sim \textit{Suzy Hits}$$

$$\textit{Bottle Shatters} = \textit{Suzy Hits} \ \vee \ \textit{Billy Hits}$$

Causal models thus encode and formalize specific token counterfactual relationships between elements in these models.

Lewisian counterfactuals are clear forerunners of causal models. The function of the latter is primarily to model subtleties of counterfactual relationships between causes and effects without Lewis’ modal realism or semantic framework. One specific respect in which Lewis set the stage for contemporary projects was his focus on demarcating *actual causation* rather than, for example, merely possible causes or mere probabilities.

Roughly, actual causation is the causal relationship between two actual events *c* and *e*. The goal of focusing on actual causation is to, in Weslake’s words, “[eliminate] all of the non-causes of an effect that can be discerned at the level of counterfactual structure.” (forthcoming, p. 1) The project of determining actual causation seeks to separate mere would-be causes, like preempted rock-throws, from actual causes. As Lewis showed, this project is harder than it looks when appealing to counterfactual structure alone.

Despite the considerable recent enthusiasm about the modelling method, some (most notably Briggs 2012) remain skeptical of the usefulness of causal models. If causal models are just Lewisian counterfactuals formalized, their extra value must stem simply from their formalization and specificity. It remains an open question whether this methodology will yield results more impressive than careful applications of the Lewisian approach.

There is an independent methodological question about whether the Lewisian approach and the causal modelling approach are attempting the same philosophical project, and whether they aspire to the same *desiderata*. One line takes the causal modelling approach to be more akin to causal *explanation* than causation itself, with distinct *desiderata* and success conditions for both projects. Others such as Hitchcock (2007) propose several distinct concepts of causation, with causal models tracking one concept and Lewisian counterfactuals tracking another concept more in line with folk intuitions. Some causal modellers, however, purposely collapse this line so that the two projects are viewed to be one and the same.

Topics with Ancestors in the Lewisian Approach

Besides giving rise to the research program of responding to and shoring up the counterfactual theory, Lewis' approach to causation gave rise to many other important related topics.

The past decades have generated a profusion of philosophical and empirical work alike on the supposed *context sensitivity* of counterfactuals. Lewis' own work generally doesn't address this feature. In *On the Plurality of Worlds* (1986), he notes that each event has an objective group of causes. Even his later influence account, within which there is room for context-sensitivity, largely focuses on objectively discernable relationships of counterfactual covariation. In contrast to Lewis' extreme realist approach, much work since homes in on contextual and pragmatic features that affect and even determine the truth conditions for counterfactuals. Schaffer (2005) argues for the *contrastivity* of causal claims: c rather than c^* causes e rather than e^* , where c and e are actual events and c^* and e^* are unactualized, contextually specified contrasts. To use Schaffer's own example to illustrate: Jane's moderate smoking rather than abstaining causes her lung cancer, but Jane's moderate smoking rather than heavy smoking does not. For Schaffer, causation is a quaternary rather than a binary relation, holding between actual events and their contrasts. Not including the contrastive events results in a failure to see the full causal picture. Menzies (2004b) similarly argued that difference-making required context-sensitivity in order to be fully true and informative.

There is now an associated body of work concerned with empirical tests of the context-sensitivity of counterfactuals. A volume from Hoerl, McCormick, and Beck (2012) includes numerous empirical studies of everyday reactions to counterfactuals. Livengood and Rose (2016) suggest that the folk do not see counterfactual dependence as necessary for causation. Such studies are important when intertwining actual linguistic practice with contextual truth conditions for counterfactuals.

The recent explosion of research on the pragmatics of counterfactuals is an outcropping of Lewis' semantic system that undergirds his analysis of causal counterfactuals. Von Fintel (2001) and Gillies (2007) use special sequences of counterfactuals, *reverse sobel sequences*, to argue against the standard Lewisian semantics for counterfactuals. The canonical example of a reverse Sobel sequence is the following:

(1a) If Sophie had gone to the parade, she would have seen Pedro dance.

(1b) But of course, if Sophie had gone to the parade and been stuck behind someone tall, she wouldn't have seen Pedro dance.

This series seems true. But read in the reverse order—(1b) to (1a)—the sequence seems false. Order seems to make a difference to the consistency of certain groups of counterfactuals. Moss (2012) defends Lewisian semantics against reverse Sobel sequences by appealing to pragmatics. More recently, Karen Lewis (forthcoming) has recently argued that both approaches are untenable.

While Lewis famously denied *backtracking* readings of counterfactuals such as “If Joanna hadn't woken up, it would have been because her alarm hadn't sounded”—there has been some recent interest in defenses of such readings. Khoo (forthcoming) argues that “counterfactuals quantify over a suitably restricted set of historical possibilities from some contextually relevant past time”, licensing backtracking readings of counterfactuals. Joyce (2010) endorses backtracking in causal reasoning. Penelope Mackie (2014) suggests that backtracking readings of counterfactuals are sometimes as natural as non-backtracking readings.

A closely associated literature questions Lewis' commitment to *temporal asymmetry*, the fixity of the passed coupled with openness of the future. In “Counterfactual Dependence and Time's Arrow,” (1979) Lewis suggests that the

direction of causal dependence matches the direction of temporal flow: effects depend on their causes because the present and future depend on the past. (He does accept the conceptual possibility of time-reversed causation, but does not focus on it.) This thesis of the *asymmetry of overdetermination*, which holds that earlier events are massively overdetermined by later events but not vice versa, is partly responsible for the huge literature on whether time is symmetric or asymmetric, and whether the arrow of causation must match the arrow of time. Albert (2000) picks up the gauntlet and argues for the time-asymmetry of counterfactual dependence, while Price (1992) famously argues against temporal asymmetry more generally.

A current surge of interest in the metaphysics and causal profile of *omissions*—events that don't occur, such as the failure to water one's plant—has roots in Lewis' (2004) "Void and Object". There he argues that causation isn't actually a relation, since a relation requires *relata*, and omissions are nothing at all. There are a cluster of currently live issues surrounding causation by omission, including what omissions are, whether and how omissions cause, and which omissions cause things if any of them do. Lewis' claim that omissions are nothing at all has been challenged by philosophers thinking variously that omissions reduce to positive events described negatively (Schaffer 2000b, 2000c, 2004, 2010), negative properties, and even possibilities (Bernstein 2014). Questions about whether or not omissions count as real causes have been batted around vigorously. Dowe (2001) defends causation by omission as "quasi-causation": something very much like causation, but not exactly worthy of the full causal honorific. Dowe proposes what he calls the "intuition of difference", the idea that omissions and negative events seem fundamentally causally different than their positive event counterparts. Schaffer (2000b, 2000c) argues forcefully against this intuition of difference, suggesting that omissive causes are ubiquitous and non-mysterious, and can be normal causes, effects, and intermediaries. Finally, a debate traceable to the Lewisian approach but begun in earnest by Menzies (2004b) calls attention to *the problem of profligate omissions*: given the counterfactual theory of causation, if any omissions count as causes, then all of them do. For example, the counterfactual "If I hadn't failed to water my plant, the plant wouldn't have died" is true, but the counterfactual "If the Queen of England hadn't failed to water my plant, the plant wouldn't have died" is also true. Options for handling this problem

include giving up the counterfactual account, using contextual or pragmatic elements to distinguish between salient omissions and non-salient ones, or including norms in the causal relation.

This latter surprising outcropping of the omissions literature-- the idea that causation is *normative*, or at least has normative dimensions—has gained force in recent years on both philosophical and empirical grounds. Thompson (2003) and McGrath (2005) argue that causation is irreducibly normative: whether *c* is a cause of *e* depends partly on norms and expectations governing whether *c* should have happened. Causal modellers have gotten in on the act, holding that norms can help us distinguish between default and deviant variables in formal models of causal relations. And recent empirical work by Henne et. al (2017) and Alicke et. al (2011) suggests that the folk concept of causation is heavily intertwined with normative concepts.

Finally, a current strain of *counterfactual skepticism* presses worries about the very possibility of truth conditions for counterfactuals that don't contain explicit probabilities. "If I had dropped the vase, it would have shattered" seems straightforwardly true. But there is an extremely small probability that the vase could have naturally reconstituted itself post-dropping. Such examples motivate DeRose's (1999) and Hajek's (MS) worry that counterfactuals can never be true. Karen Lewis (2016) argues that contextual semantics can avoid counterfactual skepticism. But metaphysicians and epistemologists alike remain concerned about the problem.

Topical Descendants of Lewis' Project

I now turn to Lewis' influence on present-day hot topics in metaphysics more generally. The advent and popularity of work on *grounding*, roughly, the dependence relationship between things that are made up and the thing or things that make those things up, has clear ancestry in the Lewisian approach to causation, causal modelling, and counterfactual dependence. Schaffer (himself a Lewis acolyte) calls grounding "something like metaphysical causation. Just as causation links the world across time, grounding links the world across levels." (2016, p. 122) Some like Bennett (2017) take grounding to be akin to causation, or even causation itself.

At first, grounding and Lewisian causation might seem to be unlikely bedfellows. But putative similarities between them have been set out as a way to flesh out the concept of grounding and make it less mysterious. Like counterfactual dependence, grounding is irreflexive, asymmetric, and intransitive. Some have postulated that the temporal symmetry of one event causing another is a kind of metaphysical priority. Others have taken up the project of giving formal models of the grounding relation akin to formal models of causal relationships. Schaffer holds that comparing metaphysical dependence to causal dependence is the best way to make sense of the otherwise mysterious notion of grounding. Karen Bennett also argues that causation might be construed as a kind of “making up” relation like the other so-called “building relations” of constitution, composition, realization, and so on. Alastair Wilson (forthcoming) argues that many unusual causal structures such as preemption and overdetermination have obvious parallels in grounding. A recent spate of literature has challenged this parallel. Bernstein (2016b) argues that grounding is not causation and not even like causation, despite their apparent similarities. Koslicki (2016) argues that the apparent similarities of formal models of both gloss over their deep differences.

Another topic locus where Lewis’ system was especially prescient was in predicting the role that *laws* would play in metaphysical theories of causation in explanation. Currently, there is debate over whether there are metaphysical laws akin to those that are seen to bind causal relationships: for example, whether and how rules govern what sorts of things can compose what other sorts of things, and in what circumstances. As laws are seen to undergird causal explanations, so, too are laws seen to undergird metaphysical explanations. Wilsch (2015) lays out a notion of metaphysical explanation in terms of metaphysical laws. Kment (2014) suggests that metaphysical laws undergird metaphysical explanations.

Finally, *impossible worlds* have become a hot topic of late, and are being utilized in several philosophical contexts. Though the modal realism central to Lewis’ causal system extends only to metaphysically possible worlds, impossible worlds have recently been invoked to account for a variety of metaphysical concepts. Nolan (2014) suggests that many metaphysical concepts are *hyperintensional*, or requiring distinctions finer-grained than mere possible worlds. For example, a poor mathematician might believe that

4=4 but disbelieve that 2+2=4. Every possible world where 2+2=4 is a world where 4=4. It is only in bringing in impossible worlds that we can account for the differences in mental content. A currently popular strand of hyperintensional metaphysics is concerned with *counterpossibles*, counterfactuals with metaphysically impossible antecedents, such as “If Hobbes had squared the circle, small children in rural India would not have cared.” There is a lively debate over whether such counterfactuals can be true or false non-vacuously. One school of thought, *vacuism*, holds that all counterpossibles are vacuous due to their impossible antecedents. If the antecedent is impossible, vacuists hold, “anything goes”. *Non-vacuists*, in contrast, argue that counterpossibles can be made sense of non-vacuously. The most compelling arguments for this conclusion invoke intuitive differences between counterpossibles: “If Hobbes had squared the circle in private, children in rural India would not have cared” seems true, while “If Hobbes had squared the circle in private, then the roses would have turned blue” seems false. Bernstein (2016a) takes non-vacuism one step farther, arguing that impossible events generate counterfactual dependencies, and thus can be causes. Finally, a related debate questions whether a similarity metric can be applied to impossible worlds, like the one Lewis applies to possible worlds. Nolan (1997) proposes a Strangeness of Impossibility condition: any possible world, no matter how bizarre, should be considered closer to actuality than any impossible world. Berto (2013) holds that it is intuitive that some impossible worlds are closer to actuality than others, while Brogaard and Salerno (2013) propose a closeness relation based on feature-sharing propositions.

Final Remarks

Lewis’ theories of causation set the stage not only for philosophical progress on the topic for the next half-century, but for fruitful outcroppings of the system across philosophical subfields and subtopics. The influence of his theory of causation, including his foundational modal realism and semantic programs, cannot be overstated. Even contemporary metaphysical topics such as grounding, laws, impossibility, and temporal asymmetry have their roots in Lewis’ seminal 1973 paper, and the various debates that followed. His influence is sure to maintain its strength for years to come.

References

- Albert, D. 2000. *Time and Chance*. Cambridge, MA: Harvard University Press.
- Alicke, M., Rose, D. and Bloom, D. 2011. "Causation, Norm Violation, and Culpable Control." *Journal of Philosophy* 108 (12):670-696.
- Bennett, K. 2017. *Making Things Up*. Oxford: Oxford University Press.
- Bernstein, S. 2014. "Omissions as possibilities" *Philosophical Studies* 167 (1):1-23.
- 2015. "A Closer Look at Trumping" *Acta Analytica* 1 (1):1-22.
- 2016a. "Omission Impossible" *Philosophical Studies* 173: 2575.
- 2016b. "Grounding Is Not Causation", *Philosophical Perspectives*, 30: 21–38.
- Berto, Francesco, "Impossible Worlds", The Stanford Encyclopedia of Philosophy (Winter 2013 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2013/entries/impossible-worlds/>](https://plato.stanford.edu/archives/win2013/entries/impossible-worlds/).
- Briggs, R. 2012. "Interventionist Counterfactuals", *Philosophical Studies*, 160 (1):139-166.
- Brogaard, B. & Salerno, J. 2013. "Remarks on Counterpossibles: Synthese 190: 639.
- Collins, J. (2007). "Counterfactuals, causation, and preemption." In D. Jacquette (Ed.), *Philosophy of logic* (pp. 1127–1144). Amsterdam: Elsevier.
- DeRose, K. 1999. "Can It Be That It Would Have Been Even Though It Might Not Have Been?" *Philosophical Perspectives* 13: 385-413.
- Dowe, P. 2001. A Counterfactual Theory of Prevention and 'Causation' by Omission. *Australasian Journal of Philosophy*, 79: 216-26.
- von Fintel, K. 1997. "Can It Be That It Would Have Been Even Though It Might Not Have Been?" In *Philosophical Perspectives 11: Mind, Causation, and World*, James Tomberlin, editor, 385–413. Blackwell Publishers, Ltd., Oxford.
- Gillies, T. 2007. "Counterfactual Scorekeeping." *Linguistics and Philosophy*, vol. 30: 329–360.
- Henne, P., Pinillos, A. & De Brigard, F. 2017. "Cause by Omission and Norm: Not Watering Plants". *Australasian Journal of Philosophy* 95 (2):270-283.
- Hajek, A. MS. "Most Counterfactuals are False".

Hempel, C. (1965): Aspects of Scientific Explanation and Other Essays in Philosophy of Science. New York: Free Press.

Hitchcock, C. 2001. "The Intransitivity of Causation Revealed in Equations and Graphs", *Journal of Philosophy*, 98: 273–99.

---- 2007. "Three Concepts of Causation". *Philosophy Compass* 2 (3):508–516.

---- 2011. "Trumping and Contrastive Causation" *Synthese* 181 (2):227–240.

Hoerl, C., McCormack, I & Beck, S.R. (eds.), 2012. *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*. Oxford: Oxford University Press.

Joyce, J. 2010. "Causal Reasoning and Backtracking" *Philosophical Studies* 147: 139.

Khoo, J. Forthcoming. "Backtracking Counterfactuals Revisited", *Mind*.

Kment, B. 2014. *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.

Koslicki, K. 2016. "Where Grounding and Causation Part Ways: Comments on Schaffer". *Philosophical Studies* 173 (1):101-112.

Lewis, D. 1973., "Causation", *Journal of Philosophy*, 70: 556–67.

---- 1979., "Counterfactual Dependence and Time's Arrow" *Noûs* 13 (4):455-476.

---- 1981., "Are We Free to Break the Laws ?", *Theoria* 47, 113-121.

---- 1986., *The Plurality of Worlds*, Oxford: Blackwell.

---- 2000., "Causation as Influence", *Journal of Philosophy*, 97: 182–97

----- 2004. "Void and Object" in John Collins, Ned Hall & L. A. Paul (eds.), *Causation and Counterfactuals*. MIT Press. pp. 277-290.

Lewis, Karen. 2016. "Elusive Counterfactuals" *Noûs*. Vol. 50 Issue 2, pp. 286-313.

----- forthcoming. "Counterfactual Discourse in Context". *Noûs*.

Livengood, J. and Rose, D. (2016). Experimental philosophy and causal attribution. In Sytsma, J. and Buckwalter, W., editors, *A Companion to Experimental Philosophy*. Blackwell.

- Mackie, J. L., 1965. "Causes and Conditions", *American Philosophical Quarterly*, 12: 245–65.
- Mackie, P. 2014. "Counterfactuals and the Fixity of the Past". *Philosophical Studies* 168 (2):1-19.
- McGrath, S. Causation by Omission: A Dilemma. 2005. *Philosophical Studies* 123:125–148.
- Menzies, P. 2004a. "Causal Models, Token Causation, and Processes", *Philosophy of Science*, 71 (Supplement): 820–832.
- 2004b. "Difference-Making in Context", in Collins, Hall, and Paul 2004, pp. 139–80.
- 2014. "Counterfactual Theories of Causation", The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual/>.
- Moss, S., 2012, *On the Pragmatics of Counterfactuals*. *Noûs*, 46: 561–586.
- Nolan, D. 1997, "Impossible Worlds: A Modest Approach", *Notre Dame Journal of Formal Logic*, 38: 535–72.
- "Hyperintensional Metaphysics" 2014. *Philosophical Studies* 171 (1):149-160.
- Pearl, J., 2000. *Causality*, Cambridge: Cambridge University Press.
- Price, H. 1992. *Mind*. "Agency and Causal Asymmetry" Vol 101, No. 403, pp. 501-520
- Schaffer, J., 2000a. "Trumping Preemption", *Journal of Philosophy*, 9: 165–81.
- 2000b Causation by Disconnection. *Philosophy of Science* 67 :285– 300.
- 2000c. 'Causes Need Not Be Physically Connected to Their Effects: The Case for Negative Causation.' *Contemporary Debates in Philosophy of Science*, C. Hitchcock (ed.) Malden MA: Blackwell, (2004) 197–216.
- 2005. "Contrastive Causation", *Philosophical Review*, 114: 297–328.
- 2012. Disconnection and Responsibility. *Legal Theory* 18 (2012): 399-435.
- 2016. "Grounding in the Image of Causation", *Philosophical Studies* 173(1): 49-100.

Scriven, M. (1962): "Explanations, Predictions and Laws," in H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, volume III, Minneapolis: University of Minnesota Press, 170–230.

Strevens, M. (2003), Against Lewis' New Theory of Causation: A Story with Three Morals. *Pacific Philosophical Quarterly*, 84: 398–412.

Thomson, Judith, 2003. "Causation: Omissions". *Philosophy and Phenomenological Research* 66 (1):81–103.

Weslake, B. forthcoming. "A Partial Theory of Actual Causation", *British Journal for Philosophy of Science*.