

Causation and the Problem of Disagreement*

Enno Fischer

Abstract

This paper presents a new argument for incorporating a distinction between default and deviant values into the formalism of causal models. The argument is based on considerations about how causal reasoners should represent disagreement over causes and it is defended against an objection that has been raised against earlier arguments for defaults.

*This is the author's original manuscript. Please cite the published version in *Philosophy of Science* 88(5), 773-783. doi:10.1086/714852.

A number of authors have argued for incorporating a distinction between default and deviant states into the formalism of causal models.¹ A central motivation for this has been the Problem of Isomorphism (Hall (2007); Halpern and Hitchcock (2015)). This problem arises from pairs of target systems that supposedly have isomorphic causal models but give rise to different judgements of actual causation. The idea is that the different judgements are explained by assumptions about particular values that variables typically or normally take on. These assumptions are taken to be captured by the default/deviant distinction.

However, more recently Blanchard and Schaffer (2017) have argued that key instances of the problem can be solved by revising one of the involved models such that it gives a more appropriate representation of the corresponding target system. They also suggest a generalization of this strategy, which I shall call the *adjust-the-model argument*. They argue that, when confronted with an instance of the Problem of Isomorphism, we should suspect that at least one of the involved models is not an appropriate representation of its target system. They also argue that defaults "come close to a free parameter in an otherwise so precise and objectively constrained formalism, which basically gives the theorist leeway to hand-write the result she wants" (192). Thus, according to them, the default/deviant distinction does more damage than good to the formalism of causal models.

In this paper I shall provide a more nuanced account of the benefits of the default/deviant distinction. I shall grant that Blanchard and Schaffer's criticism of defaults as a solution to the Problem of Isomorphism is right. However, there is another problem that is far less prominent: the Problem of Disagreement. I will show that this problem gives rise to a genuinely new argument for incorporating the default/deviant distinction.

The Problem of Disagreement has been introduced by Halpern and Hitchcock (2015). It arises from cases where agents disagree in their causal judgement even though they make the same assumptions about the underlying causal model. The Problem of Isomorphism is related to well-known examples of disagreement over what is 'the

¹In the following the term 'causal model' will refer to standard causal models (models without defaults), as introduced by Pearl (2000). Models with defaults will be called 'extended causal models' as introduced by Halpern and Hitchcock (2015).

cause' of a given effect as discussed, for example, by van Fraassen (1980). The main difference is that the Problem of Disagreement involves the explicit assumption that the disagreeing agents base their causal claims on the same underlying causal model. Halpern and Hitchcock take this to indicate that the agents' causal judgements depend not only on assumptions about causal structure but also on a distinction between default and deviant behaviour.

I will show that this argument allows two readings. First, it can be read as involving descriptive claims about how agents *do* reason about causal models in contexts where they disagree. This reading seems to be vulnerable to a version of Blanchard and Schaffer's adjust-the-model argument. If two agents disagree about judgements of actual causation, we should expect that these agents also disagree about the underlying causal model. Second, the argument can be read as involving prescriptive claims about how agents *should* reason about causes when they disagree. Here the adjust-the-model argument does not apply. I will argue that it would be wrong to require that the agents support their conflicting causal judgements with different models. Instead, I will argue, causal models should be understood as a representative tool that helps express causal claims that go beyond causal judgements that are based on potentially idiosyncratic normative presumptions. If understood in this way, they can help resolve disagreement over causes by giving a framework for disentangling normative and epistemic dimensions of disagreement. And this function can only be fulfilled if models incorporate the default/deviant distinction. I will illustrate this claim with an example that concerns the causal role of Search and Rescue missions in the Central Mediterranean with regard to increasing numbers of deaths through shipwreck in 2015 and 2016.

In section 1 I introduce Blanchard and Schaffer's arguments against defaults. In section 2 I introduce the Problem of Disagreement and I point out that Halpern and Hitchcock's way of employing it as an argument for defaults is vulnerable to a version of Blanchard and Schaffer's criticism. In section 3 I raise the question what the function of extended causal models should be in instances of disagreement. Based on the example of Search and Rescue Missions (section 4) I argue that defaults help us clarify disagreement over causes (section 5).

1 The Adjust-the-Model Argument

Blanchard and Schaffer put forward three main lines of criticism against incorporating the default/deviant distinction. First, the default/deviant distinction is unnecessary. Instances of the Problem of Isomorphism arise only because one of the involved models does not provide an appropriate representation of the underlying target system. Instead of incorporating the default/deviant distinction, we should adjust the models, adhering to generally accepted aptness constraints. These aptness constraints are rules for selecting a set of variables \mathcal{V} that constitutes the causal model and in the following we shall focus on the rule that "variables should not be allotted values that we are not willing to take seriously" (182). Blanchard and Schaffer take this aptness constraint to help us deal with cases like the gardener/queen example: some flowers would not have died if either the gardener or the Queen of England had watered them and it needs to be explained why we tend to identify only the gardener as an actual cause.²

"It is because we are willing to indulge in the fantasy of the gardener watering the flowers [...], but just can't imagine the queen stooping to the job, that we feel an asymmetry. If so then [the constraint to represent only serious possibilities]—which does independent work—was all we needed to explain the gardener/queen asymmetry. There is no apt causal model in which wiggling whether the queen waters the flowers wiggles the fate of the flowers, because there is no apt causal model that considers so ridiculous a scenario as the queen of England popping by, watering can in hand, to engage in random acts of gardening" (197).

Figure 1A gives a representation of the gardener/queen case that Blanchard and Schaffer consider to be problematic. They argue that this is not an apt model because $Q = 1$ represents a scenario that we are not willing to take seriously. Thus, they suggest eliminating variable Q , which leads to the simpler model in figure 1B, which reproduces the plausible verdict that only the gardener is an actual cause of the flowers' death.

²In the gardener/queen case the problem arises from a symmetry that is internal to the model, not from two causal models that have isomorphic structure.

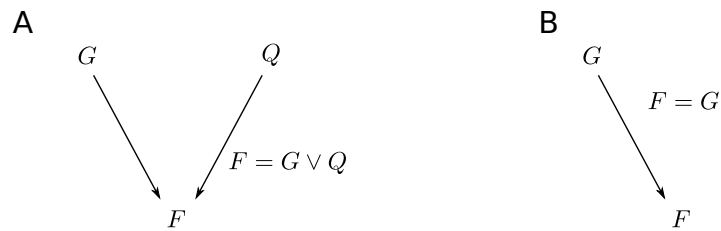


Figure 1: The flowers survive if they are watered by the gardener or by the queen.

Second, Blanchard and Schaffer argue that the default/deviant distinction involves unclarity. Most proponents of defaults relate them to an underlying theory of typicality or normality that involves a range of possibly conflicting standards. Blanchard and Schaffer worry that the unclarity associated with these notions spoils the otherwise precise theoretical framework of causal models.

Third, Blanchard and Schaffer criticise that incorporating the default/deviant distinction is psychologically implausible. Proponents of defaults assume that the judgements evoked by thought experiments like the gardener/queen case reflect judgements that arise from the competent use of a norm-laden notion of actual causation. Instead, according to Blanchard and Schaffer, the fact that causal reasoners ascribe a higher relevance to norm-violating factors or agents is to be explained by norm-related biases that interfere with the correct use of a norm-free notion of actual cause.

There is an important tension between the first line of criticism and the other two. Suppose I am a proponent of the idiosyncratic (and potentially biased) view that the queen is in charge of watering the flowers and that the gardener is not supposed to water them. According to the adjust-the-model strategy, I am supposed to represent only those scenarios that I take to be serious possibilities. Thus, I will provide a model in which variable Q is the only cause of variable F . But this is a problem. Because now my idiosyncratic view does not only spoil my judgements of actual causation, but also the corresponding causal model!

The underlying point is this. Blanchard and Schaffer argue that the default/deviant distinction is unclear and reflects biases. But they also suggest to solve cases like the gardener/queen example by adjusting the models on the basis of considerations about

what scenarios are to be taken seriously. But what is a scenario that is to be taken seriously? Presumably this depends on ideas regarding normality that are similar to those affecting the defaults—otherwise it would be easy to generate counterexamples to the strategy. But this means that the constraint on models is no less unclear. In fact, exploiting constraints on \mathcal{V} makes the problem even worse. For now the unclarity is not confined to the defaults but they infect the whole model. My argument in this paper is that there are situations where normality considerations should not affect the choice of variables in \mathcal{V} . If there are unclarity, then defaults are a better place for them.

2 The Problem of Disagreement

Consider a version of the gardener/queen case provided by Halpern and Hitchcock (2015):

"while a homeowner is on a vacation, the weather is hot and dry, her next-door neighbour does not water her flowers, and the flowers die. Had the weather been different, or had her next-door neighbour watered the flowers, they would not have died" (414f).

Halpern and Hitchcock argue that since the flowers' death depends on both the weather and the neighbour's omission it seems like a counterfactual theory of causation cannot distinguish between these factors. Yet, according to some authors (e.g. Moore (2009)) the weather is a cause of the flowers' death but not the neighbour's omission to water them because omissions generally cannot be considered to be causes. Halpern and Hitchcock flag this as the "problem of isomorphism." But, according to them, there is "an even deeper problem. There is actually a range of different opinions in the literature about whether to count the neighbour's negligence as an actual cause of the flowers' death [...]. *Prima facie*, it does not seem that any theory of actual causation can respect all of these judgments without lapsing into inconsistency" (415). This is the Problem of Disagreement.

The Problem of Disagreement arises where the following two conditions hold. First, there are two (or more) agents that have conflicting judgements of actual causation with

regard to the same target system. For example, theorists like Moore argue that only the weather is an actual cause because they think that omissions cannot be actual causes. They disagree with theorists like Lewis (2000) who think that the neighbour's negligence is also an actual cause, because they think that omissions are genuine causes. Second, it has to be the case that the opposing agents agree on the underlying causal model. In the flower case Halpern and Hitchcock take this to be a model with a graph like the one in figure 1A, and a structural equation such that the flowers die if the weather is hot and the neighbour fails to water them: $D = H \wedge \neg W$.

What exactly do the agents disagree about in such cases? According to Halpern and Hitchcock, the disagreement concerns the actual cause of the outcome. But wouldn't this imply an implausible metaphysical view according to which actual causation is subjective? Halpern and Hitchcock admit that actual causation is a subjective and context-dependent notion that is to be distinguished from an underlying and objective notion of causal structure. Yet such a notion has an important function because it indicates targets of intervention that are particularly suited from the pragmatic perspective of the agent (Hitchcock and Knobe, 2009).

Let us see how Halpern and Hitchcock account for this case of disagreement. The idea is that the default/deviant distinction gives rise to a normality ordering over the worlds that can be represented by the model. Actual causes are those factors that fulfil the Halpern-Pearl (2005) definition of actual causation plus a normality criterion. The normality criterion requires that the possible world that is needed to show that the effect depends on the cause be at least as normal as the actual world. Halpern and Hitchcock argue that "[t]hose who maintain that omissions are never causes can be understood as having a normality ranking where absences or omissions are more typical than positive events" and Halpern and Hitchcock take this to reflect "a certain metaphysical view: there is a fundamental distinction between positive events and mere absences, and in the context of causal attribution, absences are always considered typical for candidate causes" (437f). This assumption of typicality gives rise to the judgement that the only actual cause of the flowers' death is the weather.

An advocate of the view that omissions are always causes can be understood as

subscribing to an alternative normality ordering. Here the worlds in which the flowers do not die are equally normal and they are taken to be at least as normal as the world where the flowers die. Consequently, both the weather and the neighbour's negligence fulfil the normality criterion and qualify as actual causes of the flowers' death.

There are two problems with this reconstruction of the disagreement. First, it seems implausible that Moore would agree that absences are generally more normal than positive events. In fact, according to each of the many dimensions of normality, there seem to be clear counterexamples. Living humans more frequently breathe than not, functional smoke detectors remain silent (unless there is smoke), we are legally and morally required to help those whose lives are in danger. The kind of metaphysical point that Beebe and Moore make with regard to the causal status of omissions is independent of claims regarding the normality of omissions. Thus, it seems Halpern and Hitchcock have chosen an example where defaults do not do the explanatory work that they expect them to do.

Second, suppose for the sake of the argument that there is an agent who believes that absences are always considered typical and, thus, never can be causes. Moreover, suppose that the agent complies to the constraint that causal models should only represent scenarios that are to be taken seriously. According to the agent's beliefs, it is a very far-fetched possibility that omissions like the one of the neighbour are causes. Thus, the rules of appropriate modelling command that she leave out the variable representing the neighbour's negligence. But if this is the case, then this agent disagrees with the proponent of absences as causes already at the level of the standard causal models.

So, if we take the Problem of Disagreement to give rise to an argument for defaults, it seems like this argument faces the same difficulties as the argument from the Problem of Isomorphism. In particular, there is not really a problem in the first place if we choose what seem to be the most plausible representations of the agents' beliefs. The claim that there are agents who disagree about actual causes but agree on the underlying causal model seems to involve implausible empirical assumptions about the involved agents' sets of beliefs.

3 What is the Function of Extended Causal Models?

In the remainder of the paper I will argue that there is an alternative reading of the Problem of Disagreement that involves prescriptive claims about how disagreeing agents *should* use extended causal models. I will argue that the alternative reading shows that in some cases the default/deviant distinction is a useful extension.

What is the function of extended causal models? Halpern and Hitchcock "envision a kind of conceptual division of labour where the causal model [...] represents the objective patterns of dependence that could in principle be tested by intervening on the system, and [the normality ordering] represents the various normative and contextual factors that also influence judgments of actual causation" (2015, 435). So, it looks like causal reasoning involves considerations that are located at two distinct levels. First, there is the level of standard causal models. These represent the objective patterns of counterfactual dependence. Second, there is the level of judgements of actual causation. These judgements are influenced by the normality ordering which reflects normative and contextual considerations.

However, the conceptual division of labour does not seem to work as straightforwardly. First, objectivity on the level of standard causal models means that "once a suitable set of variables has been chosen, there is an objectively correct set of structural equations among those variables" (431f). Thus, the causal model itself is not objective. For the choice of the set of variables (and their possible values) is likely to be governed by criteria that are sensitive to normative and contextual factors as well (such as in Blanchard and Schaffer's treatment of the gardener/queen case). Second, even the judgements of actual causation need to have some objective core. Otherwise they could hardly help us "identify appropriate targets of corrective intervention" (432).

If causal models (plus information about the variables' actual values) and claims of actual causation are so similar, couldn't we just make do with one of them? No. Claims of actual causation are highly selective. And this has the advantage that they can guide agency very straightforwardly by indicating the best targets of intervention. Presumably, causal models are somewhat closer to the objective structure because

they allow representing larger chunks of it. They express complex counterfactual dependencies that are not captured by a simple claim of the form 'X = x is an actual cause of Y = y.' These larger chunks still depend upon norms, but do so to a lesser degree, for selection does not have to be constrained so narrowly.

In the following we shall see that the Problem of Disagreement helps to indicate one distinctive advantage of causal models, understood along these lines: they can help us provide a representation of disagreement of causes that is more conducive to resolving the disagreement than the bare claims of actual causation. Moreover, I shall argue that this function is sometimes (but not always) crucially facilitated by incorporating the default/deviant distinction.

4 An Example: Search and Rescue Missions

According to Frontex,³ the European Border Control Agency, Non-governmental Search and Rescue missions (NGO SARs) are an actual cause of the increase of the number of deaths in the Central Mediterranean in the period from 2015 to 2016. On the other hand, it has been argued that NGO SARs are only one factor acting within a complex causal structure, and that it is erroneous to describe NGO SARs as *the* cause of the increase. I will look at a study performed by Forensic Oceanography⁴ and show that the most natural way to understand their criticism of Frontex's claim is to see it as an attack on Frontex's assumptions about the causal model.

Let us begin with a closer look at the claims put forward in the Frontex report. The report describes an increase of the number of deaths of refugees and states that

"it transpired that both border surveillance and SAR missions close to, or within, the 12-mile territorial waters of Libya have unintended consequences. Namely, they influence smugglers' planning and act as a pull factor that compounds the difficulties inherent in border control and saving lives at sea. Dangerous crossings on unseaworthy and overloaded vessels were organised

³The following is based on the risk analysis report for 2017 (FRONTEX (2017)).

⁴Forensic Oceanography is part of the Forensic Architecture agency located at Goldsmiths, University of London.

with the main purpose of being detected by EUNAVFOR Med/Frontex and NGO vessels. Apparently, all parties involved in SAR operations in the Central Mediterranean unintentionally help criminals achieve their objectives at minimum cost, strengthen their business model by increasing the chances of success. Migrants and refugees – encouraged by the stories of those who had successfully made it in the past – attempt the dangerous crossing since they are aware of and rely on humanitarian assistance to reach the EU" (FRONTEX, 2017, 32).

Thus, the presence of SARs (both NGO and state-led operations) near the Libyan coastline is said to give a sense of security that encourages migrants and refugees to risk their lives. This has two effects. First, smugglers can offer crossings that are more risky. Second, there is an overall increase in attempted crossings.

The report also states that "[c]losely related issues are the safety of migrants and refugees and, most significantly, the increasing number of fatalities" (32). After reporting estimates of the fatalities in 2016 the report states that "[t]he increasing number of migrant deaths, despite the enhanced EUNAVFOR Med/Frontex surveillance and NGO rescue efforts, seems paradoxical at first glance" (33). But then the report relates the increase of fatalities to a change in the smugglers' tactics: "[t]he rising death toll mainly results from criminal activities aimed at making profit through the provision of smuggling services at any cost" (33).

It seems fair to assume that the above quoted passages can be summarized by the causal model displayed in figure 2A. In the model S represents the presence of SARs, C is a factor that represents the risk level of the individual crossing and the number of attempted crossings, and D represents the number of deaths. It is claimed that an increase in S leads to an increase in C , that an increase in C leads to an increase in D , and that an increase in S also leads to a direct decrease in D . The narrative does not allow a more detailed quantification of these functional relations. But there is a possible reading of the narrative according to which the increase of deaths via the route $\langle S, C, D \rangle$ is larger than the decrease via the route $\langle S, D \rangle$.

The Forensic Oceanography report (Heller and Pezzani, 2017) identifies the pull-

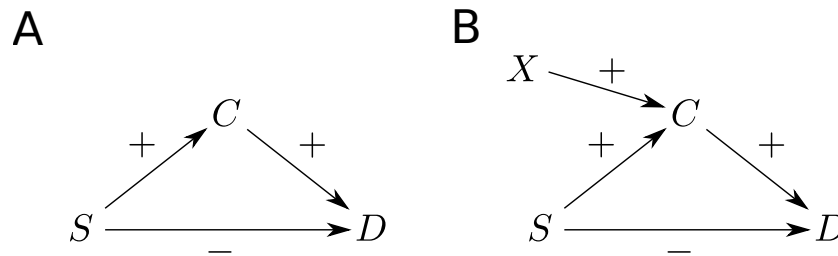


Figure 2: A simple model of the "pull-factor" narrative. A: Frontex. B: Forensic Oceanography.

factor claim as part of a toxic narrative within a "de-legitimisation and criminalisation campaign" directed at non-governmental search and rescue missions. The aim of the report is an empirical assessment of the claims put forward by Frontex. The report can be understood as challenging the structural relation between *S* and *C* as stated by Frontex and adding a variable *X* that feeds into *C* and that explains the increase in risk level and number of attempted crossings from 2015 to 2016. Relevant factors, among others, are the availability of seaworthy vessels, involvement of Libyan militia and Libyan Coast Guard.

The Forensic Oceanography report also describes non-governmental search and rescue missions as a continuation or replacement of preceding state-led search and rescue missions. In particular, the report claims that "[a]iming to deter migrants from crossing the Mediterranean, the EU and its member states pulled back from rescue at sea at the end of 2014, leading to record numbers of deaths. Non-governmental organisations (NGOs) were forced to deploy their own rescue missions in a desperate attempt to fill this gap and reduce casualties." That is, whereas the Frontex report suggests that there is a new kind of search and rescue missions that explains the increase, the Forensic Oceanography report describes the presence of search and rescue activity in the Mediterranean as a default condition.

5 The Role of Defaults

The disagreement between the Frontex report and the Forensic Oceanography report concerns the question whether the presence of SARs led to an increase in the number of

deaths. The underlying question is: why are refugees willing to risk their lives? The presence of SARs (and stories about how they guarantee safety on sea) is considered to be one factor. However, there are at least three further kinds of factors: (i) the situation in the home country, (ii) the hope for a better life in the EU, (iii) the absence of alternative pathways into the EU (legal pathways, or simply pathways that are not as risky).

Suppose each of these factors corresponds to a variable in a causal model such that a variable describing the willingness of refugees to risk their lives depends upon these variables. The disagreement about the causal role of SARs involves agents that have opposing views about which of these variables represent possible scenarios that are to be taken seriously—for functional, legal, and moral reasons. For example, there is disagreement about the moral and legal feasibility of cutting back life-saving missions on sea.

How *should* this disagreement be represented? One way would be to require that the involved agents agree on a set of variables \mathcal{V} by including all variables that are at stake in the debate and represent their disagreement on the level of the default/deviant distinction. From a humanitarian perspective, for example, life-saving missions would be the moral and legal default state. By contrast, certain opposing agents might want to describe the absence of SARs as the default state. But both kinds of agents would be required to include a variable representing SARs.

Alternatively, one could require the views to be expressed by different standard causal models that reflect the individual views about what scenarios are to be taken seriously. This is what is suggested by the adjust-the-model strategy. The advantage is that such models do not incorporate the default/deviant distinction, which is considered unclear. The disadvantage, however, is that now the unclarity occurs in a disagreement about which scenarios are to be represented by the model in the first place.

The problem with this strategy is that it leaves unclear whether agents disagree for normative or for epistemic reasons. Suppose agent A_1 does not include a particular variable X in her standard causal model even though agent A_2 thinks that X is a cause of Y . Does agent A_1 mean to say that a change in X would merely amount to a scenario that is not to be taken seriously? Or does agent A_1 mean to imply that a change in

X would not make a difference to Y ? Extended causal models fare better in this kind of context. They provide the formal resources that help the involved agents to point out where disagreement arises for normative reasons and where it arises for epistemic reasons. Agent A_1 would be required to include X into the model and clarify whether she takes Y to be independent of X or merely considers X to represent scenarios that from her particular point of view are highly abnormal.

This is particularly important in cases where it is likely that disagreement arises not only about norms but also about the underlying counterfactual dependencies. The core of Frontex's pull factor claim is the counterfactual dependency of C on S . This claim is difficult to assess directly. It involves non-trivial assumptions about the refugee's dispositions to risk their lives. It is also difficult to assess in an interventionist fashion. For performing testing interventions on the target system is unfeasible in practice. Instead Frontex supports the pull-factor claim by a comparison of the risk levels in 2015 and 2016 and relates this to an increase of the NGO SAR activity over this period. But this argument is valid only if all other potential causes for an increased risk level remain constant over this period. In Frontex's selective causal model it looks like this is the case. A more encompassing model such as the one provided by the Forensic Oceanography report, however, suggests that Frontex's claims are unwarranted. In order to warrant the pull-factor claim in the context of such a more encompassing model the Frontex report would have to show that the influence of these other factors is irrelevant.

6 Conclusion

The adjust-the-model strategy gives rise to a powerful objection to existing arguments for defaults that are based on the Problem of Isomorphism and the Problem of Disagreement. In this paper I have suggested a prescriptive reading of the Problem of Disagreement that provides a new argument for defaults that is not undermined by the adjust-the-model strategy. In cases of disagreement extended causal models should represent assumptions about the underlying causal structure that are shared by the involved agents, while the defaults should account for the normative disagreement. This helps keeping normative

disagreement apart from disagreement about the underlying counterfactual structure.

References

- Blanchard, Thomas and Schaffer, Jonathan. Cause without default. In Helen Beebe, Huw Price, Christopher Hitchcock, editor, *Making a Difference*, pages 175–214. Oxford University Press, 2017.
- FRONTEX. Risk analysis for 2017, 02 2017. URL http://frontex.europa.eu/assets/Publications/Risk_Analysis/Annual_Risk_Analysis_2017.pdf.
- Hall, Ned. Structural equations and causation. *Philosophical Studies*, 132:109–136, 2007.
- Halpern, Joseph Y. and Hitchcock, Christopher. Graded causation and defaults. *The British Journal of the Philosophy of Science*, 66:413–457, 2015.
- Halpern, Joseph Y. and Pearl, Judea. Causes and explanations: A structural-model approach. Part I: Causes. *The British Journal for the Philosophy of Science*, 56(4):843–887, 2005.
- Heller, Charles and Pezzani, Lorenzo. Blaming the rescuers, 06 2017. URL <https://blamingtherescuers.org/>.
- Hitchcock, Christopher and Knobe, Joshua. Cause and norm. *The Journal of Philosophy*, 106(11):587–612, 2009.
- Lewis, David. Causation as influence. *The Journal of Philosophy*, 97(4):182–197, 2000.
- Moore, Michael. *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford University Press, 2009.
- Pearl, Judea. *Causality. Models, Reasoning, and Inference*. Cambridge University Press, 2000.
- van Fraassen, Bas C. *The Scientific Image*. Oxford University Press, 1980.