Cognition 161 (2017) 80–93

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Normality and actual causal strength

Thomas F. Icard [a,*], Jonathan F. Kominsky [b], Joshua Knobe [c]

[a] Department of Philosophy and Symbolic Systems Program, Stanford University, United States
[b] Department of Psychology, Harvard University, United States
[c] Program in Cognitive Science and Department of Philosophy, Yale University, United States

ARTICLE INFO

Article history:
Received 1 July 2016
Revised 4 January 2017
Accepted 9 January 2017
Available online 01 February 2017

Keywords:
Causal reasoning
Normality
Counterfactuals
Actual causation
Sampling
Bayes nets

ABSTRACT

Existing research suggests that people's judgments of actual causation can be influenced by the degree to which they regard certain events as normal. We develop an explanation for this phenomenon that draws on standard tools from the literature on graphical causal models and, in particular, on the idea of probabilistic sampling. Using these tools, we propose a new measure of actual causal strength. This measure accurately captures three effects of normality on causal judgment that have been observed in existing studies. More importantly, the measure predicts a new effect ("abnormal deflation"). Two studies show that people's judgments do, in fact, show this new effect. Taken together, the patterns of people's causal judgments thereby provide support for the proposed explanation.

1. Introduction

Judgments of actual causation—concerning the extent to which a given event or factor caused some outcome—have been at the center of attention in work on causal cognition. One intriguing phenomenon that has long been recognized is that people's judgments of actual causation can be influenced by the degree to which they regard certain events as *normal*. In recent years, this effect has been explored both in experimental studies and in formal models (e.g., Halpern & Hitchcock, 2015; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Phillips, Luguri, & Knobe, 2015).

Considerable debate remains about how to explain the effect. One approach would be to posit some independent factor, outside the core processes involved in causal cognition, that explains the impact of normality. For example, one might hypothesize that the impact of normality is the result of a motivational bias or of conversational pragmatics (e.g., Alicke, Rose, & Bloom, 2011; Driver, 2008; Samland & Waldmann, 2016). Our aim is to explore a different approach. We suggest that the impact of normality might be explained by basic facts about how causal cognition works. Our explanation will rely on standard tools familiar from the literature on graphical causal models and, in particular, on a specific way of thinking about computations over these causal models involving probabilistic sampling. Drawing on these ideas, we propose a *measure of actual causal strength*. Our hypothesis is that this actual causal strength measure will help to explain the impact of normality.

The key evidence for this hypothesis comes from facts about the precise pattern of the impact of normality on causal judgment. The most well-studied effect in this domain is the tendency whereby people are inclined to regard abnormal events as more causal in certain cases. However, as we will see, the actual pattern is considerably more complex. There are also cases in which people's judgments about the causal status of a given event depend on the normality of *other* events, and these effects in turn depend on the details of the causal structure in question (Section 2). It can be shown that the causal strength measure proposed below accurately captures the details of these patterns (Section 4). More importantly, this measure generates a novel prediction, namely, that there should be cases in which abnormal events are systematically regarded as *less* causal. Two new experiments show that this prediction is in fact borne out (Section 5). Taken together, the patterns thereby provide support for the present approach.

2. Three effects of normality on actual causation judgments

Before discussing the impact of normality on people's actual causation judgments, it may helpful to clarify the notion of normality itself. To begin with, we need to distinguish two kinds of norms. First, there are purely *statistical norms*. For example, winter months in Oregon generally tend to be cloudy and overcast, so if Oregon ever had a sunny winter, this weather could be said to be violating a statistical norm. Second, there are *prescriptive norms*. These norms are constituted not by purely statistical tendencies but by the way things ought to be or are supposed to be. Suppose we believe that the police ought to accord criminal defendants cer-

* Corresponding author.
E-mail address: icard@stanford.edu (T.F. Icard).

http://dx.doi.org/10.1016/j.cognition.2017.01.010
0010-0277/

tain rights. Even if we do not believe that the police actually do tend to accord defendants these rights, we might think that failing to do so is a violation of a prescriptive norm.

A question arises as to which of these two types of norms are reflected in ordinary judgments of actual causation. As explained below, existing research suggests that actual causation judgments are influenced by both kinds of norms. More strikingly, these two kinds of norms show the same pattern of impact on such judgments. As a result, researchers have suggested that it might be helpful to posit a single undifferentiated notion of normality that integrates both statistical and prescriptive considerations (Halpern & Hitchcock, 2015; Kominsky et al., 2015). On this approach, an event counts as "abnormal" to the extent that it either violates a statistical norm or violates a prescriptive norm, and as "normal" to the extent that it follows both of these types of norms. Difficult questions arise about precisely how statistical and prescriptive considerations are integrated into an undifferentiated notion, but we will not be resolving those questions here (cf. Bear & Knobe, in press). Instead, we focus on three specific ways in which normality—both statistical and prescriptive—impacts people's intuitions about actual causation.

2.1. First effect: abnormal inflation

Abnormal inflation is the simplest of the three effects. We will eventually be introducing a formal framework in which it can be described more precisely, but for the moment, we offer the following rough characterization:

> *Suppose that an outcome depends on a causal factor C as well as an alternative causal factor A, such that the outcome will only occur if both C and A occur. Then people will be more inclined to say that C caused the outcome when they regard C as abnormal than when they regard C as normal.*

This basic effect appears to arise both for statistical norms and for prescriptive norms.

It has been known for decades that actual causation judgments can be influenced by statistical norms (Hilton & Slugoski, 1986). Suppose that a person leaves a lit match on the ground and thereby starts a forest fire. In such a case, the fire would not have begun if there had been no oxygen in the atmosphere, and yet we would not ordinarily say that the oxygen caused the fire. Why is this? The answer appears to involve the fact that it is so (statistically) normal for the atmosphere to contain oxygen. Our intuitions should therefore be very different if we consider a case in which the presence of oxygen is abnormal. (Suppose that matches were struck on a regular basis but there is never a fire except on the very rare occasions when oxygen is present.) In such a case, people should be more inclined to regard the presence of oxygen as a cause.

Strikingly, this same effect arises for prescriptive norms. Consider the following case:

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take pens, but faculty members are supposed to buy their own.
>
> The administrative assistants typically do take the pens. Unfortunately, so do the faculty members. The receptionist has repeatedly e-mailed them reminders that only administrators are allowed to take the pens.
>
> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later, that day, the receptionist needs to take an important message... but she has a problem. There are no pens left on her desk.

Faced with this case, participants tend to say that the professor caused the problem (Knobe & Fraser, 2008; Phillips et al., 2015). But now suppose that we change the first paragraph of the case in such a way as to make the professor's action not violate a prescriptive norm:

> The receptionist in the philosophy department keeps her desk stocked with pens. Both the administrative assistants and the faculty members are allowed to take the pens, and both the administrative assistants and the faculty members typically do take the pens. The receptionist has repeatedly e-mailed them reminders that both administrators and professors are allowed to take the pens.

Faced with this latter version, participants are significantly less inclined to say that the professor caused the problem (Phillips et al., 2015). Yet the two cases do not appear to differ from the perspective of statistical normality; the difference is rather in the degree to which the agent violates a prescriptive norm. The result thereby suggests that prescriptive norms impact causal judgments.

Within existing work, this first effect has been investigated in far more detail than the others we will discuss (see, e.g., Danks, Rose, & Machery, 2014; Phillips et al., 2015; Samland, Josephs, Waldmann, & Rakoczy, 2016). One of the most important findings to come out of this work is that the effect really does involve prescriptive considerations and cannot be reduced to a matter of purely statistical norms. First, one can explicitly pit the prescriptive against the statistical. In one study, participants were told that administrative assistants were allowed to take pens and faculty members were not (a prescriptive norm) but that in actual fact administrators never did take pens while faculty members always did (a statistical norm). People's judgments ended up being affected more by the prescriptive than by the statistical, with participants tending on the whole to say that the administrative assistant did not cause the problem while the faculty member did (Roxborough & Cumby, 2009). Second, one can look at cases in which different people have different prescriptive judgments. For example, one study looked at controversial political issues (abortion, euthanasia) and found that people who had opposing moral judgments about these issues arrived at correspondingly opposing causal judgments about people who performed the relevant actions (Cushman, Knobe, & Sinnott-Armstrong, 2008).

Yet, though existing work clearly shows that both statistical and prescriptive norms can lead to abnormal inflation, controversy remains regarding the explanation of this effect. Researchers have suggested that the effect might arise as a result of conversational pragmatics (Driver, 2008), motivational bias (Alicke et al., 2011), relativity to frameworks (Strevens, 2013), responsibility attributions (Sytsma, Livengood, & Rose, 2012), or people's understanding of the question (Samland & Waldmann, 2016). Here, we will be exploring a general approach that has been defended by a number of researchers in recent years, namely, that abnormal inflation reflects a process in which certain counterfactuals are treated as in some way more relevant than others (Blanchard & Schaffer, 2016; Halpern & Hitchcock, 2015; Knobe, 2010; Phillips et al., 2015).

If one simply looks at the abnormal inflation effect in isolation, it seems that any of these theories might be able to predict the experimental findings. The advantage of the account we will be offering emerges most clearly when we broaden the scope of our inquiry, looking at a number of different effects and trying to develop an account that predicts the pattern as a whole.

2.2. Second effect: supersession

Supersession is an effect whereby the apparent normality of one factor can actually influence the degree to which *other* factors are regarded as causes. The effect can be characterized roughly as follows:

> *Suppose an outcome depends on a causal factor C as well as an alternative causal factor A, such that the outcome will only occur*

if both C and A occur. Then people will be less inclined to say that C caused the outcome if A is abnormal than if A is normal.

In other words, it is not just that a given factor is regarded as more causal when it is abnormal; a factor will be also be regarded as more causal when other factors are normal. This effect also arises for both statistical and prescriptive norms.

Turning first to the case of statistical norms, consider the following scenario:

> Alex is playing a board game. Every turn of the game involves simultaneously rolling two six-sided dice and flipping a fair coin. Alex will either win or lose the game on his next turn.
>
> Alex will only win the game if the total of his dice rolls is greater than 2 AND the coin comes up heads. It is very likely that he will roll higher than 2, and the coin has equal odds of coming up heads or tails.
>
> Alex flips the coin and rolls his dice at exactly the same time. The coin comes up heads, and he rolls a 12, so just as expected, he rolled greater than 2. Alex wins the game.

Now contrast that with a case in which the second paragraph is slightly modified:

> Alex will only win the game if the total of his dice rolls is greater than 11 AND the coin comes up heads. It is very unlikely that he will roll higher than 11, but the coin has equal odds of coming up heads or tails.

The difference between these two cases is solely in the normality of the dice roll. The success of the dice roll is statistically normal in the first case, statistically abnormal in the second. Yet this difference actually leads to a change in the degree to which people regard the coin flip as a cause: participants were significantly less inclined to say that Alex won because of the coin flip when the dice roll was abnormal than when it was normal (Kominsky et al., 2015).

This same effect then arises for prescriptive norms. In one study, participants were asked to imagine a motion detector that goes off whenever two people are in the room at the same time. Suzy and Billy enter the room at the same time, and the motion detector goes off. In one condition, Billy is supposed to be in the room, while in the other condition he is specifically not supposed to be in the room. Suzy was judged to be significantly less a cause of the motion detector going off when Billy violated the prescriptive norm than when he acted in accordance with the prescriptive norm (Kominsky et al., 2015).

2.3. Third effect: no supersession with disjunction

The supersession effect arises in cases where the causal structure is conjunctive. That is, it arises in cases where there are two distinct factors such that the effect will only occur if *both* factors are present. However, turning to cases in which the causal structure is disjunctive, i.e., cases in which the effect will occur if *either* factor is present, we find a quite different pattern:

Suppose an outcome depends on a causal factor C as well as an alternative causal factor A, such that it will only occur if either C or A occurs. Then people are just as inclined to say that C caused the outcome when A is abnormal as they are when A is normal.

Existing studies have put this claim to the test by comparing disjunctive cases to conjunctive cases and looking for an interaction whereby manipulations of normality do not have the impact in disjunctive cases that they do in conjunctive ones. This interaction arises both for statistical norms and for prescriptive norms.

For statistical norms, we can see the effect by looking at the case of the coin flip and dice roll described above. One can simply modify the rules described in that case so that Alex wins if he succeeds either on the coin flip or on the dice roll. When the rules are changed in this way, the supersession effect disappears. Participants are just as inclined to see the coin flip as causal when the dice roll is abnormal as they are when it is normal (Kominsky et al., 2015).

Precisely the same result then arises for the prescriptive norm case with the motion detector. When participants are told that the motion detector will go off if at least one person is in the room, the supersession effect again disappears. Participants are just as inclined to see Suzy as the cause when Billy's act violates a prescriptive norm as when it does not (Kominsky et al., 2015).

2.4. Summary

Across three different effects, prescriptive norms appear to have the same qualitative impact as statistical norms. If there had only been an impact from one type of norm or the other, the obvious approach would have been to explain the impact in terms of that type of norm in particular. Indeed, even after these effects were demonstrated across both kinds of norms, some researchers have argued that we should still attempt to explain the impact of each separately (Samland & Waldmann, 2016). However, this plurality of explanations becomes unnecessary if we can find a single account that explains the impacts of both types of norms, especially if it would further extend to any other categories of norms that have not been as well explored (e.g., non-moral prescriptive norms of "proper functioning," cf. Lagnado & Gerstenberg, 2015).

3. Causal models and strength measures

To explain these effects, we will be offering a *measure of causal strength*. That is, we will be offering a formal measure of the degree to which one event caused another.

Within the existing literature, work on measures of causal strength has focused primarily on capturing the impact of purely statistical considerations. Accordingly, we begin by introducing a formal framework familiar from the literature on causal Bayesian networks, and we use this framework to characterize the effects of normality as they arise for purely statistical norms. We then discuss measures of causal strength that have been proposed in the existing literature. As we will see, none of these measures can account for the three effects even when we restrict attention to only statistical normality.

In the following section we will introduce a type of measure that does capture the three effects in cases involving statistical norms, and we will explain how this measure naturally explains the prescriptive cases as well.

3.1. Causal Bayes nets

Let $\mathcal{G}$ be a finite directed acyclic graph with vertices V, and let $\mathcal{X} = \{X_v\}_{v \in V}$ be a set of random variables indexed to V with joint probability distribution $P(\mathcal{X})$. We say a pair $\mathcal{N} = \langle \mathcal{G}, P \rangle$ is a Bayesian network, or Bayes net, if P can be factored in the following way:

$$P(\mathcal{X}) = \prod_{v \in V} P(X_v | \mathrm{pa}_{X_v})$$

where pa_{X_v} denotes the set $\{X_{v'} : v' \text{ is a parent of } v \text{ in } \mathcal{G}\}$. This captures the assumption that each variable is independent of its non-descendents conditional on its parents, which means that the only parameters of a Bayes net are these conditional distributions $P(X_v | \mathrm{pa}_{X_v})$.

The natural interpretation of Bayes nets is causal: we generally include a link from one variable to another if the first has a direct causal influence on the second. In fact, the idea that people rely on representations very much like Bayes nets has been shown consistent with a wide array of data on causal learning and inference in children and adults (for a review, see Sloman & Lagnado, 2015). One of the key ideas is that Bayes nets can be used not just for ordinary probabilistic inferences such as conditionalization based on obser-

vations, but also for distinctively causal manipulations known as *interventions* (Pearl, 2009; Spirtes, Glymour, & Scheines, 1993; Woodward, 2003). Our account will make use of this framework.

Intervening on a Bayes net $\mathcal{N} = \langle \mathcal{G}, P \rangle$ involves setting some variable X to a specific value x. This gives rise to a new, "manipulated" Bayes net $\mathcal{N}_{X=x} = \langle \mathcal{G}_{X=x}, P_{X=x} \rangle$, where in the graph $\mathcal{G}_{X=x}$ we cut all links to the node representing X, so that is has no parents, and in $P_{X=x}$ we have $P_{X=x}(X = x) = 1$, leaving all other conditional distributions as before. We can then use this to infer what *would* happen under various suppositions. We adopt standard notation for interventions (Pearl, 2009). Thus, given a network $\mathcal{N} = \langle \mathcal{G}, P \rangle$ we will write

$$P(Y|do(X = x)) \stackrel{\text{def}}{=} P_{X=x}(Y)$$

only obliquely referring to the manipulated network $\mathcal{N}_{X=x}$. We will also use somewhat nonstandard notation to refer to a negative intervention:

$$P(Y|do(X \neq x)) \stackrel{\text{def}}{=} P_{X\neq x}(Y),$$

where $P_{X\neq x}$ is just like P, except that $P_{X\neq x}(X = x) = 0$ and $P_{X\neq x}(X = x')$ is the renormalized probability of $X = x'$, i.e., $\frac{1}{Z}P(X = x')$, with $Z = \sum_{x' \neq x} P(X = x')$.

3.2. Desiderata

We will assume that the way people represent the motivating examples from Section 2 can be (at least to a first approximation) captured by a simple 3-node graph with two possible causes, C and A, and one effect, E (known in the literature as an "unshielded collider"), as depicted in Fig. 1. Assuming the random variables A, C, E are all binary—taking on values 0 and 1—and given a distribution P that factors over this graph, we are interested in two central cases, where the effect is some deterministic function of the causes:

- Conjunctive: $P(E|C, A) = \min(C, A)$
- Disjunctive: $P(E|C, A) = \max(C, A)$

In other words, the conjunctive version has E on (value 1) if both C and A are on, off (value 0) otherwise. This kind of model would describe the scenario with the pens, for example: the receptionist has a problem ($E = 1$) just in case both the administrator takes a pen ($C = 1$) and the professor takes a pen ($A = 1$). By contrast, the disjunctive version has E on if at least one of C or A is on. This describes the disjunctive scenarios from Section 2: e.g., the motion detector goes off ($E = 1$) just in case either Billy enters the room ($A = 1$) or Suzy enters the room ($C = 1$).

Given this setup, we can specify the desiderata we would like a causal strength measure to satisfy, corresponding to the three effects discussed in the previous section. Suppose we have a functional $\kappa_P(C, E)$ that measures the strength of cause $C = 1$ on effect $E = 1$ under distribution P. Given two distributions P_1 and P_2, defined such that P_2 represents a case in which one of the causal variables involves a norm violation and P_1 represents a case in which both are normative (or more normative than in P_2), we should expect different patterns depending on whether these distributions correspond to conjunctive or disjunctive scenarios:

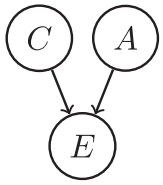


Fig. 1. Simple unshielded collider structure.

Abnormal Inflation: In the conjunctive case, suppose $P_1(C) > P_2(C)$ but that $P_1(A) = P_2(A)$. Then $\kappa_{P_1}(C, E) < \kappa_{P_2}(C, E)$.

Supersession: Again in the conjunctive case, suppose instead that $P_1(C) = P_2(C)$ but that $P_1(A) > P_2(A)$. Then $\kappa_{P_1}(C, E) > \kappa_{P_2}(C, E)$.

No Supersession with Disjunction: In the disjunctive case, suppose again that $P_1(C) = P_2(C)$ and $P_1(A) > P_2(A)$. Then, nonetheless, $\kappa_{P_1}(C, E) = \kappa_{P_2}(C, E)$.

The assumption we are making is that, in the statistical norm violation cases, the probabilities of C and A are being manipulated. If a person's representation of the situation appropriately reflects this assumption, then these three patterns indeed capture the desiderata motivated by the three effects. We would like to find such a measure κ that would satisfy all of them.

3.3. Causal strength measures

A number of causal strength measures have been discussed in philosophy and psychology (for a recent survey, see Fitelson & Hitchcock, 2011). Many of these were not intended to capture people's ordinary causal intuitions but rather to characterize the true nature of causal strength. Moreover, many of them have been considered in the context of general causation—assessing the extent to which a given variable is causally related to another variable, generally speaking—rather than actual causation. Some of these measures were also intended for application only in restricted settings (e.g., models with disjunctive causes). Nonetheless, it is worth seeing why these existing proposals do not automatically capture the patterns discussed so far, even the patterns as they pertain solely to statistical normality.

One prominent proposal is aimed at capturing the intuition that a cause C should raise the probability of the effect E above its unconditional value (Reichenbach, 1956; Spellman, 1997; Suppes, 1970) (we adopt the name *SP* following Mandel, 2003, who attributes the measure to Spellman, 1997; *SP* also invokes Suppes, 1970):

$$SP: \quad P(E|C) - P(E)$$

Note that *SP* depends on $P(C)$, the prior probability of C, which some authors have argued to be inappropriate for a measure of the causal relation between C and E (see especially Cheng & Novick, 1992). A number of researchers have therefore proposed a slight variation, known as ΔP, that is independent of $P(C)$ (e.g., Cheng & Novick, 1992; Jenkins & Ward, 1965)[1]:

$$\Delta P: \quad P(E|C) - P(E| \sim C)$$

If causal strength is measured according to *SP* or ΔP, then whenever $P(E| \sim C)$ is very high, causal strength will always be low. For this and other reasons, Cheng (1997) proposed the so called power-PC account:

$$\text{power-PC}: \quad \Delta P / P(\sim E| \sim C)$$

An even more elaborate measure, in a sense generalizing both ΔP and power-PC, has been suggested by Pearl (2009). His measure is designed to capture the extent to which a cause C is both necessary and sufficient for E, where these correspond to the following counterfactuals:

- *Necessity*: If C were not to occur, E would also not occur.
- *Sufficiency*: If C were to occur, E would also occur.

Pearl assesses the joint probability of these counterfactual statements and derives the following measure, which he calls the *probability of necessity and sufficiency* (*PNS*):

[1] As a notational abbreviation we sometimes write, e.g., $P(C)$ instead of $P(C = 1)$ and $P(\sim C)$ instead of $P(C = 0)$.

$$PNS: \quad P(C=E=1)P(E=0|do(C=0),E=C=1)+ \\ P(C=E=0)P(E=1|do(C=1),E=C=0)$$

Under certain assumptions (in fact shared by the unshielded collider structure), *PNS* is equivalent to ΔP. Nevertheless, we mention *PNS* because it is similar in spirit to our own proposal, in that we will also be invoking notions of necessity and sufficiency.

The causal judgments we are considering in this paper involve scenarios in which the existence of certain causal relations is explicitly stipulated. In many empirical studies, judgments (especially of general causal strength) are instead elicited after presentation of contingency data, either sequentially or in summary form.[2] In this context, Griffiths and Tenenbaum (2005) have argued that a model of causal strength ought to incorporate aspects of structure induction, inferring whether a causal link between C and E exists at all. They show that existing measures such as ΔP and POWER-PC can be naturally interpreted as already assuming that a link does exist and inferring parameters of a corresponding graphical model. Their *causal support* model is designed specifically for this setting involving contingency data, and is given by the log likelihood ratio of the data with and without the assumption that a link exists between C and E. Because the "data" in our scenarios are simply statements that a causal relation exists, this model does not readily apply to the cases we are considering, so we will not consider it further here.

3.4. Assessing the desiderata

How do these different strength measures fare with respect to our desiderata? We summarize the predictions in the particular cases of interest for all of these measures in Table 1. To derive these predictions we simply use the expressions given above in Section 3.3, plugging in the relevant values. For instance, in the conjunctive model we assume $P(E|C) = P(A)$, and so on. Some of these measures were not intended to be used in arbitrary settings. For example, Cheng (1997) is explicit that POWER-PC is intended for cases of non-interacting disjunctive causes. Nonetheless, we find it useful to understand what predictions these measures would make if we used them in these settings.

All of these measures manifest supersession in the conjunctive case. However, only POWER-PC shows no supersession in the disjunctive case, and only *SP* shows abnormal inflation. (Notably, a fourth effect called "abnormal deflation", to be presented and analyzed below in Section 5, is predicted by none of these measures.) Indeed, we have found no measure (including all of those surveyed by Fitelson & Hitchcock, 2011) that satisfies more than two of the desiderata; and as these examples demonstrate, they often differ in which desiderata they satisfy. This is not even to mention the parallel effects of non-statistical, prescriptive normality.

It should be emphasized once more that this is not intended to be a criticism of any of these models, in as far as they are designed for other purposes (e.g., characterizing the true nature of causal strength, predicting people's judgments about general causal strength in disjunctive scenarios, etc.). The intention is rather to highlight the need for a new kind of measure capturing judgments of actual causal strength, which can account for these characteristic patterns involving different varieties of normality.

4. A new actual causal strength measure

Each of the measures discussed in the previous section was motivated by an intuition about the nature of causal strength. We now present a new type of measure motivated by a different kind of intuition. This measure is not motivated by an intuition about what causal strength really is (e.g., how it would make sense to characterize causal strength if we were using that notion as part of a systematic scientific inquiry). Instead, it is motivated by an intuition about the psychological processes people go through when they are making causal judgments.

Table 1
Predicted causal judgments of C under existing causal strength measures.

	Conjunctive	Disjunctive
SP	$(1-P(C))P(A)$	$(1-P(C))(1-P(A))$
ΔP/*PNS*	$P(A)$	$1-P(A)$
POWER-PC	$P(A)$	1

More specifically, we will be drawing on the idea that people make causal judgments through a process of *probabilistic sampling* (e.g., Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014). We introduce an algorithm that proceeds by sampling counterfactuals. We then show that this algorithm converges to a particular measure. The hypothesis is that this measure accurately captures the patterns in the impact of normality on people's judgments of actual causation.

4.1. Counterfactual analysis of actual cause

A long line of work in philosophy and psychology maintains that in judging whether some actual event C caused E, people consider counterfactual scenarios involving C and E (Lewis, 1973). Though there are many ways of spelling out this counterfactual approach (see, e.g., Hall, 2004; Halpern & Pearl, 2005; Lewis, 2000, among many others), it is generally agreed that, at a minimum, one must check whether the cause C was in some sense *necessary* for effect E to happen. For actual causes, this corresponds roughly to the following counterfactual:

> *Actual Necessity*: If C had not occurred, E also would not have occured.

Dual to necessity, a number of researchers have proposed that judgments of actual causation also involve a notion of *robust sufficiency* (Hitchcock, 2012; Lombrozo, 2010; Woodward, 2006):

> *Robust Sufficiency*: Given that C occurred, E would have occurred even if background conditions had been slightly different.

Importantly, existing research suggests a difference between the role of necessity and the role of sufficiency (Woodward, 2006). When it comes to necessity, it is generally enough if a causal factor just happens to be necessary given various highly contingent facts about the actual situation. Even if slight changes in background conditions would have made that causal factor unnecessary, this will not count much against its causal strength. By contrast, when it comes to sufficiency, the relevant claim has to hold across a variety of possible background conditions. It is not enough if the background conditions just happened to come out in such a way that the causal factor ended up being sufficient; the factor will be seen as fully causal to the extent that it would remain sufficient even if the background conditions were somewhat different.

Thus, we suggest that a quantitative measure of actual strength might in some way combine actual necessity and robust sufficiency. A question now arises as to how people assess each of these factors and how they combine the two.

4.2. Sampling propensity and normality

One possible way to answer this question would be to pick out a specific set of counterfactuals and then claim that people's causal cognition proceeds by deterministically evaluating all and only the counterfactuals in this set. An alternative approach would be to suggest that people's causal cognition relies on a process of *probabilistic sampling*. On this latter approach, the idea would be that a wide variety of counterfactuals could in principle be relevant and people show a certain probability of evaluating each of them.

[2] There are in fact interesting questions about how the patterns discussed above manifest themselves in such settings (see, e.g., Danks et al., 2014).

The idea that many cognitive processes can be understood in terms of probabilistic sampling algorithms has garnered much excitement and attention in recent years (see Griffiths, Vul, & Sanborn, 2012 for a review, and Icard, 2016 for discussion related to the present proposal). Typically, the distributions from which samples are drawn correspond to normative models that make sense from a statistical point of view. Specifically, sampling algorithms in this literature are usually offered as a tractable means the mind might use to approximate difficult inductive computations. For instance, instead of trying directly to compute a complex integral $\int P(X)\,\phi(X)\,dX$, such as an expected utility value, it might make sense to draw some number K of sample values $X^{(k)}$ from $P(X)$ and approximate the integral by means of the sample average $\frac{1}{K}\sum_{k=1}^{K}\phi(X^{(k)})$. These matters are closely tied to statistical questions, e.g., questions involving prediction. However, when one shifts to thinking in terms of sampling propensities, a new possibility suggests itself. The probability of sampling a given counterfactual need not be proportional simply to its purely statistical properties. Instead, sampling propensities might be proportional to overall *normality*. In other words, sampling propensities may be impacted by a blend of statistical and prescriptive considerations.

Existing research provides some support for this hypothesis. When participants are given a vignette and asked to provide a counterfactual, they are more likely to mention possibilities they regard as statistically frequent but also possibilities they regard as prescriptively good (Kahneman & Miller, 1986; McCloy & Byrne, 2000). In addition, when participants are given a counterfactual and asked to rate the degree to which it is relevant or worth considering, they are more inclined to rate a possibility as relevant to the extent that it conforms to prescriptive norms (Phillips et al., 2015). These findings provide at least some initial evidence in favor of the claim that people are drawn to consider possibilities that do not violate prescriptive norms.

Returning to the specific context of the unshielded collider causal structure, making this assumption means that, for example, $P(C)$ will be higher not only when C is statistically more likely, but also when C is assumed to be prescriptively more normal. That is, people will be inclined to imagine C scenarios, rather than non-C scenarios, both to the extent that C is likely and to the extent that C is prescriptively normal. Difficult questions arise when we inquire into the exact nature of this amalgamation of statistical and prescriptive normality (see, e.g., Bear & Knobe, in press). Without answering these questions, we will assume only that $P(C)$ does positively correlate with an undifferentiated notion of normality. Significantly, this assumption allows the characterization of the desiderata above in Section 3.2 to remain unchanged.

Thus, to motivate our measure of causal strength we will be assuming that the parameters of the unshielded collider network—in particular the priors $P(C)$ and $P(A)$ on causes C and A—correspond with sampling propensities, which are in turn proportional to normality. However, even if the hypothesis about sampling propensity is ultimately rejected, the measure we will derive in the next section can still be maintained by replacing $P(C)$ and $P(A)$ with some measure of normality of C and A. Apart from this change, everything in the account would remain the same.

Note that we will not be directly testing the assumption that people probabilistically sample counterfactual scenarios to make causal judgments. Rather, this general picture is the inspiration for our new causal measure. To the extent that the measure captures important aspects of people's judgments, it also shows that people's judgments are compatible with this probabilistic sampling account, but further work would be required to establish that these judgments are genuinely the result of a sampling process.

4.3. Measuring necessity and sufficiency strength

Before defining our causal strength measure, we need to say something about how we understand the notions of actual necessity and robust sufficiency. Instead of settling on specific formalizations of these notions, we will suggest general ways of characterizing them, which nonetheless issue in specific proposals for necessity and sufficiency strength in the context of the disjunctive and conjunctive unshielded collider structures.

Consider first actual necessity. A prevalent assumption in the literature is that, to determine actual necessity of $X = x$ for $Y = y$, we must identify a path—what is often called an *active path*—from X to Y, and freeze all other variables $\vec{Z}$ outside the path to specific values $\vec{z}$, e.g., the values that they held in the observed situation. We then perform another intervention to set $X \neq x$ and check whether nonetheless $Y = y$. There are various proposals in the literature for how to choose $\vec{Z}$ and $\vec{z}$ (see, e.g., Halpern & Pearl, 2005; Hitchcock, 2001; Weslake, in press; Woodward, 2003). Having made such a choice, we would then assess the necessity strength of $X = x$ on $Y = y$ to be something like $P\left(Y \neq y | do(X \neq x, \vec{Z} = \vec{z})\right)$.

Because in this paper we are primarily interested in the simple unshielded collider structure, we will remain neutral on what the right general account for selecting $\vec{Z} = \vec{z}$ is, though ultimately we would of course like to have a theory of it. In our specific case, we assume this involves setting C to 0, holding fixed $A = 1$, and checking whether this is enough to make $E = 0$. Thus, in the conjunctive model, the necessity strength of $C = 1$ is just $P(E = 0|do(C = 0, A = 1)) = 1$. In the disjunctive model, the necessity strengths are both 0 if in fact the other was present. We thus assume there is some measure of actual necessity strength given by a probability, $P^{v}_{X\neq x}(Y \neq y) = P(Y \neq y| \ldots do(X \neq x) \ldots)$, which takes on these values in the cases of interest.

Consider now the concept of robust sufficiency. Here again, researchers begin with a rough, intuitive notion, and there have been a variety of different proposals about how to spell it out more precisely. Still, it seems that just about any plausible account would generate the same predictions in the unshielded collider structures that we are using as test cases. In the disjunctive case, the presence of C completely guarantees E in a way that is independent of the value of any other variable. Thus, we assume that in the disjunctive case C should have sufficiency strength 1. Similarly, in the conjunctive case, C will bring about E if and only if A is present. So we will assume that in the conjunctive case, C should have the sufficiency strength $P(A)$.

There are a number of specific accounts that would make exactly these predictions. For example, a particularly simple proposal would be to intervene to set $X = x$, resample all other variables, and then determine whether $Y = y$, thereby giving $P(Y = y|do(X = x))$ as the measure of strength of $X = x$ on $Y = y$. Alternatively, one could use Cheng's POWER-PC as a measure of sufficiency strength, as Pearl essentially does for his PNS measure.[3] In structures like the unshielded collider, this is equal to $P(Y = 1|do(X = 1), X = Y = 0)$. We will simply assume sufficiency strength is measured by some probability $P^{\sigma}_{X=x}(Y) = P(Y| \ldots do(X = x) \ldots)$, and leave the details of the general case of this measure for future work.

Table 2 summarizes what we assume about these measures of actual necessity and robust sufficiency in the specific context of the unshielded collider network. Note that in the deterministic setting we are studying, only sufficiency strength in the conjunctive case is intermediate between 0 and 1. A general theory of $P^{v}_{X\neq x}$ and $P^{\sigma}_{X=x}$ might allow for more intermediate values, e.g., allowing for partial necessity.

4.4. A measure of actual causal strength

We are now ready to introduce our causal strength measure, supposing we have chosen candidates for $P^{v}_{X\neq x}$ and $P^{\sigma}_{X=x}$. For this section let us assume for simplicity that X and Y are binary variables.

[3] Thanks to Christopher Hitchcock for this suggestion.

Table 2
Necessity/sufficiency strength of C in the unshielded collider when $C = A = E = 1$.

	Disjunctive	Conjunctive
Necessity strength $P^{v}_{C=0}(E=0)$	0	1
Sufficiency strength $P^{\sigma}_{C=1}(E=1)$	1	$P(A)$

Then we can define an elementary algorithm, inspired by the sampling view, for determining causal strength of $X = 1$ on $Y = 1$.

```
Initialize N = 0, and for k ⩽ K:
  Sample a value X^(k) from P.
  If X^(k) = 0, sample Y^(k) from P^v_{X=0}. Let N = N + (1 − Y^(k)).
  If X^(k) = 1, sample Y^(k) from P^σ_{X=1}. Let N = N + Y^(k).
Return N/K.
```

Intuitively, we simulate an X-situation and, depending on the value we sample for X, we either test for actual necessity or robust sufficiency by simulating a Y-situation. This algorithm, which is very natural from the general perspective sketched in the previous section (Section 4.2), immediately suggests a definition of actual causal strength. It is easy to see that as $K \to \infty$, the ratio N/K converges (with probability 1) to the following, which we take to be our causal strength measure:

$$\kappa_P(X,Y) \stackrel{\text{def}}{=} P(X=0)\,P^{v}_{X=0}(Y=0) + P(X=1)\,P^{\sigma}_{X=1}(Y=1) \qquad (1)$$

In words, the causal strength of $X = 1$ is simply the weighted sum of its necessity strength and sufficiency strength, these being weighted by the probability that $X = 0$ and $X = 1$, respectively.

It is now straightforward to show that this actual causal strength measure κ satisfies the three desiderata described above, assuming the values for necessity and sufficiency in Table 2. Consider first the disjunctive case when $C = A = E = 1$:

$$\begin{aligned}\kappa_P(C,E) &= P(\sim C)\cdot P(\sim E|do(\sim C,A)) + P(C)\cdot P(E|do(C))\\ &= P(C)\end{aligned} \qquad (2)$$

This makes it completely clear why No Supersession with Disjunction holds, because $\kappa_P(C,E)$ does not depend on $P(A)$ at all. Consider next the conjunctive case:

$$\kappa_P(C,E) = P(C)\,P(A) - P(C) + 1 \qquad (3)$$

Again, this allows Supersession to fall out immediately, as this expression is monotonic in $P(A)$.

Finally, let us consider Abnormal Inflation, again in the conjunctive case, and let us assume that $P(A) < 1$. (Otherwise, C should always have maximal causal strength). As $P(C)$ increases, this leads to a decrease in the second term that outweighs the increase in the first term, giving an overall decrease in $\kappa_P(C,E)$. Put intuitively: in conjunctive cases necessity strength is 1, while sufficiency strength is $P(A) < 1$. As C becomes more normal, thinking more about C puts more weight on sufficiency strength, which leads to an overall decrease in causal strength.

5. Abnormal deflation

So far, we have shown that κ can explain results from existing studies. We now turn to a new prediction of the model and report two new studies designed to test it.

As we noted above, numerous existing studies of conjunctive scenarios find an effect of Abnormal Inflation whereby people regard the more abnormal factors as more causal. Particularly in the case of prescriptive norms, there are several explanations for this effect. This effect is predicted by our model, but it is also predicted by a wide variety of other views which invoke everything from conversational pragmatics to motivational bias (Alicke et al., 2011; Driver, 2008; Samland & Waldmann, 2016; Strevens, 2013; Sytsma et al., 2012), in addition to being predicted by one of the measures discussed above (namely *SP*). However, our model predicts a further effect, for both prescriptive and statistical norms. In disjunctive scenarios, it predicts that we should see exactly the opposite pattern: people should actually regard abnormal factors as *less* causal. We will refer to this second effect as Abnormal Deflation.

The basic prediction is as follows:

> *Suppose an outcome E depends on a causal factor C as well as an alternative causal factor A, such that E will only occur if either C or A occurs. Then people will be more inclined to say that C caused E if C is normal than if C is abnormal.*

Or, in the notation from Section 4, we should have:

> Abnormal Deflation: In the disjunctive case, suppose $P_1(C) > P_2(C)$ and in fact $A = C = 1$. Then $\kappa_{P_1}(C,E) > \kappa_{P_2}(C,E)$.

This can be easily seen from Eq. (2). The intuition is again similar. In disjunctive cases C is not at all necessary, while sufficiency strength is 1. Thus, increasing the normality of C shifts the weight onto sufficiency strength, which leads to an overall increase in causal strength. The prediction that norm violators should be judged *less* causal in disjunctive cases is not considered in any previous work (to our knowledge), and therefore makes for a strong test of our model.

To test this prediction, we simply took the stimuli and procedure that had been used previously to demonstrate causal supersession and no supersession in disjunctive cases, but rather than asking for causal judgments about the causal agent whose actions are held constant across these manipulations of normative status and causal structure (the "fixed" agent), we asked about the agent who does or does not violate a norm (the "varied" agent). In conjunctive cases, we should see the well-established phenomenon of abnormal inflation, such that the varied agents are more causal when they violate a norm than when they do not. In disjunctive cases, however, the model makes the novel prediction that we should see abnormal deflation, such that the varied agents are *less* causal when they violate a norm than when they do not. We test this prediction both for violations of prescriptive norms (Experiment 1) and for violations of statistical norms (Experiment 2).

5.1. Experiment 1

Our first study tested the prediction for prescriptive norms using not only the same procedure but also some of the same stimuli from earlier experiments on causal supersession. We start with prescriptive norms because there are strong predictions from previous work that we should not find this effect: any explanation of these effects based on motivational bias or conversational pragmatics (e.g. Alicke et al., 2011; Driver, 2008; Samland & Waldmann, 2016; Strevens, 2013; Sytsma et al., 2012) would predict either abnormal inflation regardless of causal structure, or no effect. To support the broader prediction that our model should be quite general, we used several different specific scenarios, which vary on parameters that the model considers irrelevant (causal domain, social context, etc.). Therefore, we used four different sets of vignettes, one of which comes directly from previous work on causal superseding (Kominsky et al., 2015, Experiment 3), and manipulated both causal structure and whether the actions of the varied agent violated a prescriptive norm.

5.1.1. Methods

Participants. We recruited 480 workers from Amazon Mechanical Turk, restricting our sample by location (USA only) and overall work acceptance rate ($\geqslant$90%). Workers were paid $0.25 for approximately 1–2 min of work.

Materials and procedure. Each participant read one of four versions of one of four different vignettes. Each vignette had two components that could be manipulated: Causal structure (conjunctive

vs. disjunctive) and whether the varied agent violated a prescriptive norm. The overall design of the experiment was therefore a 2 (causal structure) × 2 (norm violation) × 4 (vignette) design, run fully between-subjects.

One of the vignettes was the exact one used in Experiment 3 of Kominsky et al. (2015). The other three were created specifically for this experiment, and covered a wide variety of prescriptive norm violations (ignoring a signal, violating company policy, stealing from a friend) and consequences (causing a bridge collapse, deleting information from a network, starting a car). Every vignette involved characters named Billy and Suzy. Suzy was the fixed agent, and her actions were always normative. Billy was the varied agent, and his actions were either normative or in violation of a prescriptive norm, depending on condition. An example of four variations of one of the vignettes is presented in Table 3 (reproduced from Table 3 in Kominsky et al., 2015). The other three vignettes can be found in Appendix A.

After reading the vignette, participants were asked to rate on a 1 (strongly disagree) to 7 (strongly agree) scale how much they agreed with the statement "Billy caused [the outcome]", where the outcome was whatever was appropriate to that vignette (for example, for the vignette presented in Table 3, the statement was "Billy caused the motion detector to go off").

After making this rating, participants were asked two multiple-choice manipulation check questions. The questions were unique to each vignette, but one question always validated the prescriptive norm manipulation (e.g., "Who was supposed to show up at 9 am?") and the other validated the causal structure manipulation (e.g., "The motion detector goes off when it detected how many people?"). Participants were excluded if they answered either question incorrectly.

5.1.2. Results

We excluded 54 participants prior to analysis for failing to answer the manipulation check questions correctly, leaving data from 426 participants for analysis. Mean responses for each condition are displayed in Fig. 2a. (Raw data for both experiments are available on the Open Science Framework, osf.io/j23gr.)

We analyzed agreement ratings with a 2 (causal structure) × 2 (norm violation) × 4 (vignette) ANOVA. The analysis revealed significant main effects of causal structure, $F(1,410) = 17.04$, $p < .001$, $\eta_p^2 = .040$, norm violation, $F(1,410) = 13.39$, $p < .001$, $\eta_p^2 = .031$, and vignette, $F(3,410) = 6.41$, $p < .001$, $\eta_p^2 = .045$. Critically, there was a significant interaction between causal structure and norm violation, $F(1,410) = 71.01$, $p < .001$, $\eta_p^2 = .148$, and no three-way interaction, $F(3,410) = 1.92$, $p = .125$. We provide a summary of the means by condition and vignette in Appendix A for interested readers.

To further explore the interaction between causal structure and norm violation, we conducted separate 2 (norm violation) × 4 (vignette) ANOVAs for each causal structure (conjunctive and disjunctive). In the conjunctive scenarios, we replicated the well-established abnormal inflation effect, as participants gave higher agreement ratings when Billy violated a norm ($M = 5.61$, $SD = 1.79$) than when he did not ($M = 3.37, SD = 2.11$), $F(1,211) = 77.18$, $p < .001$, $\eta_p^2 = .268$. There was a significant main effect of vignette, $F(3,211) = 2.71$, $p = .046$, $\eta_p^2 = .037$, and, unexpectedly, a significant interaction, $F(3,211) = 6.15$, $p < .001$, $\eta_p^2 = .080$, suggesting that the abnormal inflation effect was stronger in some vignettes and weaker in others.

In disjunctive scenarios, this analysis revealed abnormal deflation: participants gave *lower* agreement ratings when Billy violated a norm ($M = 3.25, SD = 2.05$) than when he did not ($M = 4.18$, $SD = 1.93$), $F(1,199) = 10.75$, $p = .001$, $\eta_p^2 = .051$. There was also a main effect of vignette, $F(3,199) = 6.75$, $p < .001$, $\eta_p^2 = .092$, but no interaction between vignette and norm violation, $F(3,199) = .42$, $p = .736$, indicating that the magnitude of the abnormal deflation effect did not differ significantly between vignettes.

5.1.3. Discussion

The impact of prescriptive norms was examined both in conjunctive cases and in disjunctive cases. For conjunctive cases, we replicated the well-established finding that when an agent does something wrong, she is regarded as more causal. However, for disjunctive cases, we found exactly the opposite pattern. When an agent did something wrong, that agent was actually regarded as *less* causal.

Within the conjunctive cases, we also observed a significant interaction such that the size of the abnormal inflation effect varied from one scenario to the next. An inspection of the means indicated that this interaction arose primarily because the abnormal inflation effect was smaller in the one scenario that involved a good outcome (Battery). In other words, participants regarded the agent who did something wrong as more causal in all conjunctive cases, but this effect was smaller in the one case where the outcome was itself good. This same interaction pattern has been observed in previous studies (Alicke et al., 2011), and it appears to be a real phenomenon. Perhaps it arises because, in addition to the effect we have been exploring here, there is also an effect of motivated cognition such that participants are reluctant to attribute good outcomes to morally bad agents.

That said, it is worth noting the overall pattern of the results. In all cases, violations of prescriptive norms led to higher causal ratings when the structure was conjunctive and lower causal ratings when the structure was disjunctive.

Table 3
Example vignette from Experiment 1 (reproduced from Kominsky et al., 2015). Each participant saw one combination of causal structure and morality.

1a) Morally good: Suzy and Billy are working on a project that is very important for our nation's security. The boss tells them both: "Be sure that you are here at exactly 9am. It is absolutely essential that you arrive at that time."	*1b) Morally bad:* Suzy and Billy are working on a project that is very important for our nation's security. The boss tells Suzy: "Be sure that you are here at exactly 9am. It is absolutely essential that you arrive at that time." Then he tells Billy:"Be sure that you do not come in at all tomorrow morning. It is absolutely essential that you not appear at that time."
2) Event: Both Billy and Suzy arrive at 9am.	
3a) Conjunctive: As it happens, there was a motion detector installed in the room where they arrived. The motion detector was set up to be triggered if *more than one person* appeared in the room at the same time. So the motion detector went off.	*3b) Disjunctive:* As it happens, there was a motion detector installed in the room where they arrived. The motion detector was set up to be triggered if *at least one person* appeared in the room. So the motion detector went off.

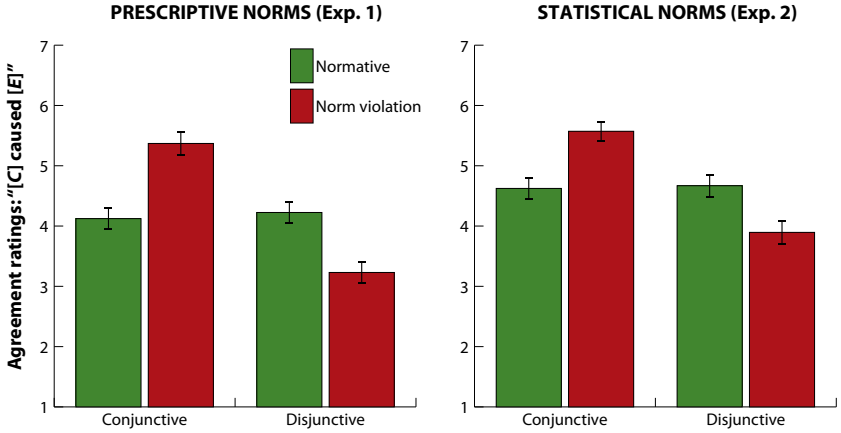


Fig. 2. Mean agreement ratings by condition for Experiments 1 and 2. Error bars represent ±1 SE mean.

5.2. Experiment 2

Experiment 2 was designed to extend the results of Experiment 1 to the case of statistical norm violation. Our approach does not distinguish between different types of norms. Indeed, it specifically holds that all norm violations should have roughly the same impact on causal judgments. So, in cases where one cause is compatible with statistical norms and the other is in violation of those norms, we should once again find both abnormal inflation in conjunctive cases and abnormal deflation in disjunctive cases.

5.2.1. Methods

Participants. We recruited 480 workers from Amazon Mechanical Turk who did not participate in Experiment 1 (exclusions based on worker ID), with the same restrictions and compensation.

Materials and procedure. We used the same 2 (causal structure) × 2 (norm violation) × 4 (vignette) fully between-subjects design of Experiment 1, with four new vignettes that involved statistical norms rather than prescriptive norms. An example vignette is provided in Table 4, and the others can be found in Appendix B. This experiment was otherwise identical to Experiment 1.

5.2.2. Results

We excluded 110 participants for answering one or both of the manipulation check questions incorrectly, leaving data from 370 participants for analysis. The results are summarized in Fig. 2b.

We analyzed agreement ratings with a 2 (causal structure) × 2 (norm violation) × 4 (vignette) ANOVA. This analysis revealed a significant main effect of causal structure, $F(1, 354) = 19.98$, $p < .001$, $\eta_p^2 = .053$, but no main effect of norm violation, $F(1, 354) = .23$, $p = .633$, or vignette $F(3, 354) = 2.11$, $p = .098$. Critically, there was a significant interaction between causal structure and norm violation, $F(1, 354) = 22.20$, $p < .001$, $\eta_p^2 = .059$, and no three-way interaction, $F(3, 354) = 1.57$, $p = .197$. We provide the means by condition and vignette in Appendix B for interested readers.

We once again explored this interaction between causal structure and norm violation with separate 2 (norm violation) × 4 (vignette) ANOVAs for each causal structure. In the conjunctive condition, we found clear abnormal inflation, with participants giving higher agreement ratings when a cause violated a norm ($M = 5.58, SD = 1.41$) than when it did not ($M = 4.58, SD = 1.88$), $F(1, 196) = 17.23$, $p < .001$, $\eta_p^2 = .081$. There was a marginal main effect of vignette, $F(3, 196) = 2.63$, $p = .051$, $\eta_p^2 = .039$, and no interaction, $F(3, 196) = 2.17$, $p = .092$, indicating that abnormal inflation occurred to roughly the same degree across all four vignettes.

In the disjunctive condition we found significant abnormal deflation, with participants giving lower agreement ratings to a cause when it violated a probabilistic norm ($M = 3.89, SD = 1.97$) than when it did not ($M = 4.67, SD = 1.73$), $F(1, 158) = 7.01$, $p = .009$, $\eta_p^2 = .04$. In addition, there was no main effect of vignette, $F(3, 158) = .36$, $p = .785$, and no interaction, $F(3, 158) = .47$, $p = .702$, indicating that the magnitude of the abnormal deflation effect did not differ significantly between vignettes.

5.2.3. Discussion

An abnormal deflation effect was found not only for prescriptive norms but also for statistical norms. The results showed the same basic pattern for both types of norms. In conjunctive scenarios, participants regarded an event as more causal when it violated a norm. By contrast, in disjunctive scenarios, participants actually regarded an event as *less* causal when it violated a norm. The present results thereby suggest that the much-studied abnormal inflation effect is in fact just one facet of a broader pattern which also involves a reversal in disjunctive cases.

Importantly, this effect can be explained using the approach we explore here but would be difficult to explain on alternative approaches. As we noted at the outset, the present account can be seen as one way of working out the details of a broader approach that has been pursued by a number of researchers. The core of this broader approach is the idea of explaining the impact of prescriptive norms in terms of people's tendency to treat normal possibilities as in some way more relevant (Blanchard & Schaffer, 2016; Halpern & Hitchcock, 2015; Knobe, 2010; Phillips et al., 2015). The fact that the present account can easily explain the deflation effect provides at least some evidence in favor of this broader approach. Subsequent research could ask whether the effect can also be explained by other accounts within the same broad family (perhaps drawing on the theoretical frameworks introduced by Halpern & Hitchcock, 2015, or by McCloy & Byrne, 2002).

By contrast, another approach would be to suggest that the impact of prescriptive norms arises because people's causal judgments are influenced by judgments that particular agents are *blameworthy*. This broad approach has been worked out in sophisticated detail by a number of different researchers, and existing work has led to the development of a variety of accounts that differ from each other in important respects (Alicke et al., 2011; Samland & Waldmann, 2016; Sytsma et al., 2012). Still, it seems that any account in this second broad family would have trouble explaining the deflation effect. Early work focused on conjunctive cases, and in those cases, it does seem plausible that the agent who violates a prescriptive norm will be regarded as more blameworthy and, for that reason, as more causal. But the deflation effect is that in disjunctive cases, the agent who violates a prescriptive norm will be regarded as *less* causal. In other words, the very agent who is doing something more morally bad will be seen as less of a cause. It is difficult to see how to explain such an effect on the assumption that people's causal judgments are being driven in some way by blame.

Of course, the present results do not thereby show that this second approach is fundamentally mistaken. It may well be that there are different processes at work in these judgments, with some effects being explained as we have proposed and others being explained in terms of blame or responsibility judgment. Still, whatever else may

Table 4
Example vignette from Experiment 2. Each participant saw one combination of causal structure and norm violation.

1) Background: Prof. Smith works at a large university. At this university, in order to get new computers from the university, faculty like Prof. Smith must send an application to two administrative committees, the IT committee and the department budget committee.	
2a) Conjunctive: Prof. Smith will be able to get her new computers if the IT committee approves her application AND the department budget committee approves her application. Both committees must approve the application for her to get the new computers.	*2b) Disjunctive:* Prof. Smith will be able to get her new computers if the IT committee approves her application OR the department budget committee approves her application. Only one of the committees needs to approve her application for her to get the new computers.
3a) Likely: The IT committee almost always approves these applications. The department budget committee also almost always approves these applications. The budget committee is notorious for approving almost every application they receive. Prof. Smith sends in her applications. Each committee meets independently and they decide without talking to each other, but their meetings are scheduled for the exact same time. The IT committee approves her application, and as expected, the department budget committee approves her application. So, Prof. Smith got her new computers.	*3b) Unlikely:* The IT committee almost always approves these applications. The department budget committee almost never approves these applications. The budget committee is notorious for turning down almost every application they receive. Prof. Smith sends in her applications. Each committee meets independently and they decide without talking to each other, but their meetings are scheduled for the exact same time. The IT committee approves her application, and surprisingly, the department budget committee approves her application. So, Prof. Smith got her new computers.

be going on, the present results do suggest that our approach is tapping into an important aspect of people's causal cognition.

6. General discussion

Existing work suggests that the impact of normality on causal judgment shows certain complex patterns. To explain these patterns, we developed a measure of actual causal strength. This measure then predicted a new effect, which we call abnormal deflation. Two experiments indicate that the abnormal deflation effect does, in fact, arise. Overall, then, existing results appear to lend at least some preliminary support to this actual causal strength measure.

One noteworthy feature of the present account is that it does not invoke processes that are external to causal cognition (blame attribution, motivation, conversational pragmatics). Rather, it explains the impact of normality in terms of certain basic facts about the way people's causal cognition itself works. Thus, to the extent that the patterns of people's causal judgments are accurately captured by this account, these patterns provide evidence for the more general view that the impact of normality is to be explained in terms of basic facts about the workings of causal cognition.

In the remainder of this section, we discuss how the model fits into the study of causal cognition more broadly, including how our model might fit with more complex causal setups.

6.1. Broader causal setups

As summarized in Table 5, the model we have presented makes very specific qualitative predictions about how a change in either of the two parameters, $P(A)$ or $P(C)$, will lead to a change in the causal strength $\kappa_P(C, E)$ of the focal cause C. It will either be monotonically related $(+)$, anti-monotonically related $(-)$, or it will effect no change in $\kappa_P(C, E)$ at all $(*)$.

Three of these effects—Abnormal Inflation, Supersession, and No Supersession with Disjunction—were already observed in previous work (Knobe & Fraser, 2008; Kominsky et al., 2015; Phillips et al., 2015). The model we presented in this paper was in fact conceived as a simple, intuitive proposal that would predict each of these (Icard & Knobe, 2016). As we have seen in Section 5, the remaining distinctive prediction about the three-variable network also captures a robust pattern in people's causal judgments, namely Abnormal Deflation, a pattern that no other existing causal strength measure predicts. While one could certainly ask further questions about the unshielded collider structure—e.g., what happens if we manipulate the normality of A and C simultaneously?—as Table 5 makes evident, this marks a reasonably comprehensive study of this particular type of causal setup.

Table 5
How does $\kappa_P(C, E)$ change as a function of $P(C)$ or $P(A)$?

	Disjunctive	Conjunctive
Change in $P(C)$	$+$ (Abnormal Deflation)	$-$ (Abnormal Inflation)
Change in $P(A)$	$*$ (No Supersession)	$+$ (Supersession)

The unshielded collider structure we have studied has been the focus of much work on causal reasoning. Nonetheless, further tests of our proposal could consider an expanded class of causal setups, along at least three dimensions. First, it would be natural to consider cases where the causal relationships themselves are non-deterministic, e.g., where $0 < P(E|A, C) < 1$. The measure κ would make a number of additional predictions about these cases, with necessity strength itself taking intermediate values. Second, one would like to investigate other complex functional relationships, including cases with non-binary variables. Third, and perhaps most obviously, exploring other graphical structures with varied numbers of variables and causal relationships among variables would be an important next step.

With these variations one would then be able to study other critical phenomena concerning actual causation. For instance, much of the philosophical literature on actual cause has focused on puzzling cases involving overdetermination, preemption, omission, and so on, which have motivated specific qualitative proposals about how actual causation works (Halpern & Hitchcock, 2015;

Halpern & Pearl, 2005; Hitchcock, 2001; Weslake, in press; Woodward, 2003). In order to deal with these more complex phenomena, we would need to settle on a more comprehensive hypothesis about how the subsidiary variable choices $\vec{Z}$ and $\vec{z}$ are made when assessing actual necessity.

Another significant question is whether the measure κ can be extended to handle not just generative causes but also preventative causes. A common proposal in the literature is to take a measure of generative causal strength $k^+(C,E)$ and extract from it a measure of preventative causal strength $k^-(C,E)$ in the following way (see, e.g., Fitelson & Hitchcock, 2011):

$$k^-(C,E) \stackrel{\text{def}}{=} -k^+(C,\sim E)$$

In our algorithm for deriving causal strength (Section 4.4 above), arriving at this result would be straightforward: to determine sufficiency strength of C, we would simply check whether $E=0$, and similarly to determine necessity strength we would check to see if $E=1$; and instead of adding 1 we would subtract 1 to our sum at each step where either of these occurred in the sampling process. Normality effects in prevention cases have been less studied than in generative cases, but our measure could be used to derive predictions about this setting as well.

Exploring these further questions could be aided if we had a normative theory of actual causal judgments, that is, an answer to the question of what "problem" actual causal judgments might be solving. In particular, it has been suggested that answering some of the difficult questions in this area—such as how to select and determine values of auxiliary variables $\vec{Z}$—could profitably be guided by such a characterization (Glymour et al., 2010; Woodward, 2016). Note, however, that answering this question might crucially involve reference to basic cognitive constraints and facts about how the mind works. This was already apparent in our characterization of the measure κ: the particular tradeoff between necessity and sufficiency was motivated by the idea that this would be cognitively natural and simple. Clearly there are further questions about how to describe these constraints and combine them with concrete tasks that might be involved in causal cognition. It is conceivable that future empirical work on the current proposal will proceed in tandem with further extensions and refinements motivated by normative theorizing.

6.2. Conclusion

We have proposed a measure of actual causal strength and provided some experimental evidence showing that it captures facets of lay causal judgments that previously proposed measures of causal strength do not. Future research could further explore the question as to whether the processes underlying people's causal judgments are accurately described by this measure.

At the same time, the precise measure we propose is perhaps best understood as arising from the combination of a number of distinct hypotheses, and these hypotheses can each be examined separately. One hypothesis is that actual causal judgments are some function of (actual) necessity judgments and (robust) sufficiency judgments. A second is that causal judgments are based on a process of probabilistically sampling counterfactual scenarios. A third is that the sampling propensity of a counterfactual scenario is proportional to an integrated notion of *normality* (incorporating both statistical and prescriptive norms).

Putting all of these hypotheses together, one can derive the specific actual causal strength measure presented here. Yet, these different hypotheses are clearly separable. Some may be correct even if others turn out to be mistaken. Thus, the most fruitful path for future research might be not just to examine the causal strength measure that arises when all of these hypotheses are put together but also to explore each of them independently.

Acknowledgement

We would like to thank Christopher Hitchcock, James Woodward, Jonathan Livengood, Adam Bear, and Jonathan Phillips for helpful comments on an earlier draft of this paper.

Appendix A. All vignettes and by-vignette results of Experiment 1

See Tables A1–A3 and Fig. A1.

Table A1
'Battery' vignette from Experiment 1.

1a) Conjunctive: Billy and Suzy inherited an unusual type of hybrid car that has two special car batteries called Bartlett batteries. The car won't start unless it has two Bartlett batteries. Having one battery isn't enough to start the car. When they got the car, both Bartlett batteries were missing.	*1b) Disjunctive:* Billy and Suzy inherited an unusual type of hybrid car that has two special car batteries called Bartlett batteries. The car won't start unless it has at least one Bartlett battery. Having a second Bartlett battery isn't necessary to start the car. When they got the car, both Bartlett batteries were missing.
2a) No violation: One day, Billy and Suzy are both out of the house. Billy is visiting his friend's house, and notices that his friend has a Bartlett battery. Billy asks his friend to sell the battery to him, and his friend says that he's willing to sell it for a fair price, so Billy buys the Bartlett battery from his friend.	*2b) Norm violation:* One day, Billy and Suzy are both out of the house. Billy is visiting his friend's house, and notices that his friend has a Bartlett battery. Billy asks his friend to sell the battery to him, but his friend says that he can't sell it because he needs it for his own car. Billy waits until his friend is in the bathroom, and then steals the Bartlett battery from his friend.
Meanwhile, on the other side of town, Suzy walks into an automotive parts shop and happens to notice that they have a single Bartlett battery in stock. Suzy decides to buy the Bartlett battery from the shop. When Billy and Suzy get home, they installed the two Bartlett batteries.	
1a) Conjunctive (con't): Since the car now had both Bartlett batteries, they were able to start the car.	*1b) Disjunctive (con't):* Since all the car needed was at least one Bartlett battery, they were able to start the car.

Table A2
'Train' vignette from Experiment 1.

1) Background: Billy and Suzy are freight train conductors. One day, they happen to approach an old two-way rail bridge from opposite directions at the same time. There are signals on either side of the bridge.	
2a) No violation: Billy's signal is green, so he is supposed to drive across the bridge immediately. Suzy's signal is green, so she is also supposed to drive across immediately.	*2b) Norm violation:* Billy's signal is red, so he is supposed to stop and wait. Suzy's signal is green, so she is supposed to drive across immediately.
3a) Conjunctive: Neither of them realizes that the bridge is on the verge of collapse. If they both drive their trains onto the bridge at the same time, it will collapse. Neither train is heavy enough on its own to break the bridge, but both together will be too heavy for it.	*3b) Disjunctive:* Neither of them realizes that the bridge is on the verge of collapse. If either of them drives their train onto the bridge, it will collapse. Either train is heavy enough on its own to break the bridge.
2a) No violation (con't): Billy follows his signal and drives his train onto the bridge immediately at the same time that Suzy follows her signal and drives her train onto the bridge. Both trains move onto the bridge at the same time, and at that moment the bridge collapses.	*2b) Norm violation (con't):* Billy decides to ignore his signal and drives his train onto the bridge immediately at the same time that Suzy follows her signal and drives her train onto the bridge. Both trains move onto the bridge at the same time, and at that moment the bridge collapses.

Table A3
'Computer' vignette from Experiment 1.

1a) Conjunctive: Billy and Suzy work for a company that has a central computer. If two people log in to the central computer at exactly 9:27am, some work e-mails will be immediately deleted.	*1b) Disjunctive:* Billy and Suzy work for a company that has a central computer. If anyone logs in to the central computer at exactly 9:27am, some work e-mails will be immediately deleted.
2a) No violation: In order to make sure that two people are available to answer phone calls during designated calling hours, the company issued the following official policy: Billy and Suzy are both permitted to log in to the central computer in the mornings, and neither of them are permitted to log in to the central computer in the afternoons.	*2b) Norm violation:* In order to make sure that one person is always available to answer incoming phone calls, the company issued the following official policy: Billy is the only one permitted to log in to the central computer in the afternoons, whereas Suzy is the only one permitted to log in to the central computer in the mornings. Billy is never permitted to log into the central computer in the morning.
3) Outcome: This morning at exactly 9:27am, Billy and Suzy both log into the central computer at the same time. Immediately, some work e-mails are deleted.	

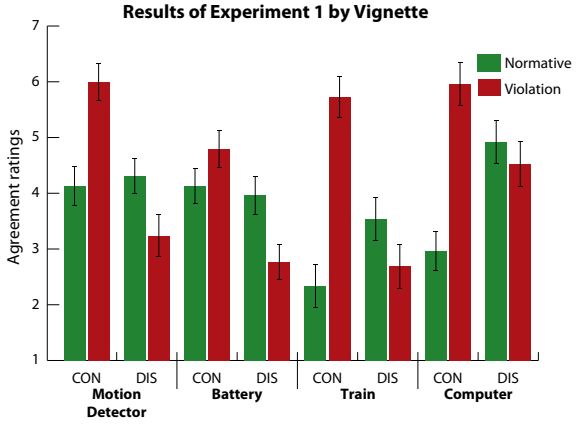


Fig. A1. Results of Experiment 1 by vignette and condition. Error bars represent ±1 SE mean.

Appendix B. All vignettes and by-vignette results of Experiment 2

See Tables B1–B3 and Fig. B1.

Table B1
'Dice' vignette from Experiment 2.

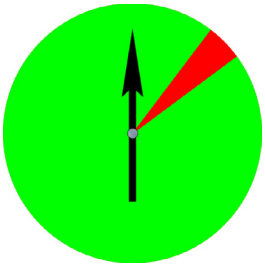
<table>
<tr><td colspan="2">1) Background: Alex is playing a board game. Every turn in the game, you simultaneously roll two six-sided dice and spin a spinner. Here is a picture of the spinner.
  </td></tr>
<tr><td>2a) Conjunctive/likely: Alex will either win or lose the game on his next turn. Alex will only win the game if the total of his dice roll is greater than 2 AND the spinner lands on green. It is very likely that he will roll higher than 2. Normally, the spinner does land on green.
Alex rolls his dice and spins the spinner at exactly the same time. He rolls a 12, so just as expected, he rolled greater than 2, and the spinner lands on green. Alex wins the game.</td><td>2b) Conjunctive/unlikely: Alex will either win or lose the game on his next turn. Alex will only win the game if the total of his dice roll is greater than 11 AND the spinner lands on green. It is very unlikely that he will roll higher than 11. Normally, the spinner does land on green.
Alex rolls his dice and spins the spinner at exactly the same time. He rolls a 12, so amazingly, he rolled greater than 11, and the spinner lands on green. Alex wins the game.</td></tr>
<tr><td>2c) Disjunctive/likely: Alex will either win or lose the game on his next turn. Alex will only win the game if the total of his dice roll is greater than 2 AND the spinner lands on green. It is very likely that he will roll higher than 2. Normally, the spinner does land on green.
Alex rolls his dice and spins the spinner at exactly the same time. He rolls a 12, so just as expected, he rolled greater than 2, and the spinner lands on green. Alex wins the game.</td><td>2d) Disjunctive/unlikely: Alex will either win or lose the game on his next turn. Alex will only win the game if the total of his dice roll is greater than 11 OR the spinner lands on green. It is very unlikely that he will roll higher than 11. Normally, the spinner does land on green.
Alex rolls his dice and spins the spinner at exactly the same time. He rolls a 12, so amazingly, he rolled greater than 11, and the spinner lands on green. Alex wins the game.</td></tr>
</table>

Table B2
'Sprinkler' vignette from Experiment 2.

1a) Conjunctive: Alex works in a building with an automatic sprinkler system and a walkway across the front lawn. If the sprinklers turn on at 8:45am AND it starts raining at 8:45am, the walkway will be flooded when Alex arrives at work.	*1b) Disjunctive:* Alex works in a building with an automatic sprinkler system and a walkway across the front lawn. If the sprinklers turn on at 8:45am OR it starts raining at 8:45am, the walkway will be flooded when Alex arrives at work.
2a) Likely: The sprinkler system is controlled by a simple timer and turns on at 8:45am every day. Because the office is in Seattle and it is winter, it usually starts raining at 8:45am every day. Winter morning rain in Seattle is an almost daily occurrence. Today, the sprinkler turned on at 8:45am and, as usual, it started raining at 8:45am. So, when Alex got to work, the walkway was flooded.	*2b) Unlikely:* The sprinkler system is controlled by a simple timer and turns on at 8:45am every day. Because the office is in New Mexico and it is summer, it almost never starts raining at 8:45am. Summer morning rain in New Mexico is very rare. Today, the sprinkler turned on at 8:45am and, unexpectedly, it started raining at 8:45am. So, when Alex got to work, the walkway was flooded.

Table B3
'Building' vignette from Experiment 1.

1) Background: In a particular building there are two businesses, a travel agency and a graphic design studio. The building's climate control system is a new design that saves energy by keeping track of the number of people in the building, and only turning on when enough people have entered the building.	
2a) Conjunctive: The climate control system will only turn on when the people who work at the travel agency AND the people who work in the design studio arrive for work. Neither office has enough employees to turn on the climate control system on their own.	*2b) Disjunctive:* The climate control system will turn on when the people who work at the travel agency OR the people who work in the design studio arrive for work. Each office has enough employees to turn on the climate control system on their own.
2a) Likely: The travel agency employees almost always arrive at 8:45am, and the design studio employees almost always arrive at 8:45am. Today, the travel agency employees arrived at 8:45am. The design studio employees also arrived at 8:45am, as usual. So, today, the climate control system turned on at 8:45am.	*2b) Unlikely:* The travel agency employees almost always arrive at 8:45am, but the design studio employees almost always arrive at 10am. Today, the travel agency employees arrived at 8:45am. Unexpectedly, the design studio employees also arrived at 8:45am to meet a deadline. So, today, the climate control system turned on at 8:45am.

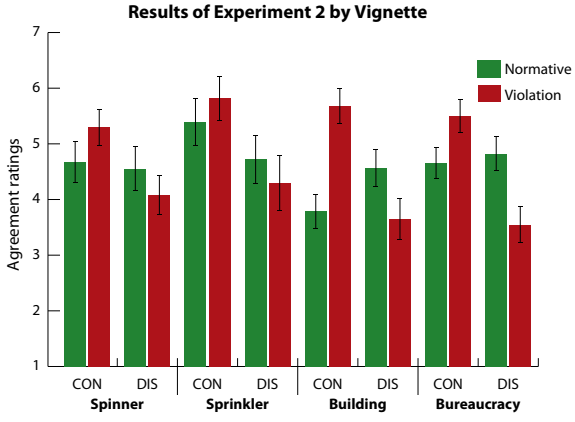


Fig. B1. Results of Experiment 2 by vignette and condition. Error bars represent ±1 SE mean.

References

Alicke, M., Rose, D., & Bloom, D. (2011). Causation, norm violation, and culpable control. *Journal of Philosophy, 108*(12), 670–696.

Bear, A., & Knobe, J. (in press). Normality: Part descriptive, part prescriptive. *Cognition*. http://dx.doi.org/10.1016/j.cognition.2016.10.024 (in press).

Blanchard, T., & Schaffer, J. (2016). Cause without default. In H. Beebee, C. Hitchcock, & H. Price (Eds.), *Making a Difference*. Oxford University Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99*, 365–382.

Cushman, F., Knobe, J., & Sinnott-Armstrong, W. (2008). Moral appraisals affect doing/allowing judgments. *Cognition, 108*(1), 281–289.

Danks, D., Rose, D., & Machery, E. (2014). Demoralizing causation. *Philosophical Studies, 171*(2), 251–277.

Driver, J. (2008). Attributions of causation and moral responsibility. *Moral Psychology, 2*, 423–440.

Fitelson, B., & Hitchcock, C. (2011). Probabilistic measures of causal strength. In P. M. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 600–627). Oxford University Press.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2014). From counterfactual simulation to causal judgment. In *Proceedings of the 36th annual meeting of the Cognitive Science Society*. .

Glymour, C., Danks, D., Glymour, B., Eberhardt, F., Ramsey, J., Scheines, R., ... Zhang, J. (2010). Actual causation: A stone soup essay. *Synthese, 175*, 169–192.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*, 334–384.

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science, 24*(4), 263–268.

Hall, E. (2004). Two concepts of causation. In J. Collins, E. Hall, & L. Paul (Eds.), *Causation and counterfactuals* (pp. 225–276). MIT Press.

Halpern, J. Y., & Hitchcock, C. (2015). Graded causation and defaults. *British Journal for Philosophy of Science, 66*(2), 413–457.

Halpern, J. Y., & Pearl, J. (2005). Causes and explanations: A structural-model approach. Part 1: Causes. *British Journal for the Philosophy of Science, 56*(4), 843–887.

Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review, 93*(1).

Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy, 98*, 273–299.

Hitchcock, C. (2012). Portable causal dependence: A tale of consilience. *Philosophy of Science, 79*(5), 942–951.

Icard, T. F. (2016). Subjective probability as sampling propensity. *The Review of Philosophy and Psychology, 7*(4), 863–903.

Icard, T. F., & Knobe, J. (2016). Causality, normality, and sampling propensity. In *Proceedings of the 38th annual conference of the Cognitive Science Society*. .

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied, 79*(1), 1–17.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 94*, 136–153.

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences, 30*(4), 315–329.

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. *Moral Psychology, 2*, 441–448.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition, 137*, 196–209.

Lagnado, D., & Gerstenberg, T. (2015). A difference-making framework for intuitive judgments of responsibility. In D. Shoemaker (Ed.). *Oxford studies in agency and responsibility* (Vol. 3, pp. 213). Oxford: Oxford University Press.

Lewis, D. (1973). Causation. *Journal of Philosophy*, 556–567.

Lewis, D. (2000). Causation as influence. *Journal of Philosophy*, 182–197.

Lombrozo, T. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology, 61*(4), 303–332.

Mandel, D. R. (2003). Judgment dissociation theory: An analysis of differences in causal, counterfactual, and covariational reasoning. *Journal of Experimental Psychology: General, 132*(3), 419–434.

McCloy, R., & Byrne, R. (2000). Counterfactual thinking and controllable events. *Memory and Cognition, 28*, 1071–1078.

McCloy, R., & Byrne, R. M. J. (2002). Semifactual "even if" thinking. *Thinking and Reasoning, 8*(1), 41–67.

Pearl, J. (2009). *Causality*. Cambridge University Press.

Phillips, J., Luguri, J. B., & Knobe, J. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities. *Cognition, 145*, 30–42.

Reichenbach, H. (1956). *The direction of time*. University of California Press.

Roxborough, C., & Cumby, J. (2009). Folk psychology concepts: Causation 1. *Philosophical Psychology, 22*(2), 205–213.

Samland, J., Josephs, M., Waldmann, M. R., & Rakoczy, H. (2016). The role of prescriptive norms and knowledge in children's and adult's causal selection. *Journal of Experimental Psychology: General, 145*(2), 125–130.

Samland, J., & Waldmann, M. (2016). How prescriptive norms influence causal inferences. *Cognition, 156*, 164–176.

Sloman, S. A., & Lagnado, D. (2015). Causality in thought. *Annual Review of Psychology, 66*(3), 1–25.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General, 126*, 323–348.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. Springer.

Strevens, M. (2013). Causality reunified. *Erkenntnis, 78*(2), 299–320.

Suppes, P. (1970). *A probabilistic theory of causality*. North Holland.

Sytsma, J., Livengood, J., & Rose, D. (2012). Two types of typicality: Rethinking the role of statistical typicality in ordinary causal attributions. *Studies in History and Philosophy of Science Part C, 43*(4), 814–820.

Weslake, B. (in press). A partial theory of actual causation. *British Journal for the Philosophy of Science* (in press).

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review*, 1–50.

Woodward, J. (2016). *Causal cognition: Physical connections, proportionality, and the role of normative theory*. Retrieved from <http://philsci-archive.pitt.edu/11630/> (Forthcoming in W. Gonzalez (Ed.) Philosophy of psychology: The conception of James Woodward).