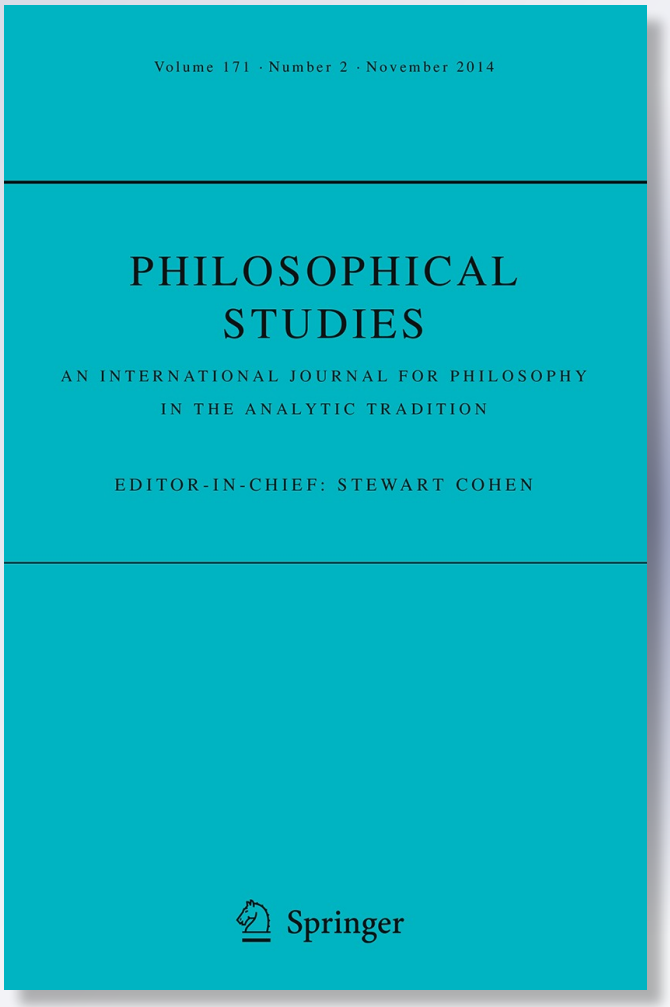
Demoralizing causation

David Danks, David Rose & Edouard Machery

Philosophical Studies
An International Journal for Philosophy in the Analytic Tradition

ISSN 0031-8116
Volume 171
Number 2

Philos Stud (2014) 171:251-277
DOI 10.1007/s11098-013-0266-8



Springer

Philos Stud (2014) 171:251–277
DOI 10.1007/s11098-013-0266-8

Demoralizing causation

David Danks · David Rose · Edouard Machery

Published online: 24 December 2013

Abstract There have recently been a number of strong claims that normative considerations, broadly construed, influence many philosophically important folk concepts and perhaps are even a constitutive component of various cognitive processes. Many such claims have been made about the influence of such factors on our folk notion of causation. In this paper, we argue that the strong claims found in the recent literature on causal cognition are overstated, as they are based on one narrow type of data about a particular type of causal cognition; the extant data do not warrant any wide-ranging conclusions about the pervasiveness of normative considerations in causal cognition. Of course, almost all empirical investigations involve some manner of ampliative inference, and so we provide novel empirical results demonstrating that there are types of causal cognition that do not seem to be influenced by moral considerations.

Keywords Causal judgment · Normative considerations · Causal reasoning · Causation · Moral judgment

D. Danks (✉)
Departments of Philosophy & Psychology, Carnegie Mellon University, 135 Baker Hall, Pittsburgh, PA 15213, USA
e-mail: ddanks@cmu.edu

D. Rose
Department of Philosophy, Rutgers University, New Brunswick, NJ, USA
e-mail: drose@philosophy.rutgers.edu

E. Machery
Department of History & Philosophy of Science, University of Pittsburgh, Pittsburgh, PA, USA
e-mail: machery@pitt.edu

1 Introduction

There have recently been a number of strong claims that normative considerations, broadly construed, influence many philosophically important folk concepts and perhaps are even a constitutive component of various cognitive processes.[1,2] For example, Knobe (2010, p. 315) writes that "[m]oral considerations actually figure in the fundamental competencies people use to make sense of the world," while Pettit and Knobe (2009, p. 602) claim:

> There is now good reason to believe there are no concepts anywhere in folk psychology that enable one to describe an agent's attitudes in a way that is entirely independent of moral considerations. The impact of moral judgments, we suspect, is utterly pervasive.

We are here concerned with versions of these claims that focus on causal judgments and the folk concept of causation. For example, Knobe (2009) argues that:

> The use of [statistical and moral] considerations is simply built into the fundamental mechanisms that subserve people's counterfactual reasoning. Any aspect of human cognition that makes use of counterfactuals will be affected in some way by the structure of these mechanisms. Since causal judgments make use of counterfactuals, and since moral considerations play a role in the mechanisms underlying counterfactual reasoning, moral considerations end up playing a role in causal judgments as well. (p. 242)

The same idea is expressed in Hitchcock and Knobe (2009): "people's causal intuitions are determined in part by judgments about the relevance of counterfactuals and (…) judgments of relevance are, in turn, determined in part by the application of norms." (p. 612) Many others have expressed similar views about the influence of normative considerations on causal judgments, though sometimes without claiming that normative considerations are a constitutive component of the cognitive processes underwriting such judgments (e.g., Alicke et al. 2011; Driver 2008a, b; Feinberg 1970; Hart and Honoré 1985; Knobe and Fraser 2008; Lombrozo 2010; McGrath 2005; Sytsma et al. 2012).

Strong claims demand strong evidence or strong arguments. While the extant empirical results may show that normative considerations sometimes influence some kinds of causal judgments, we argue that the strong claims found in the recent literature on causal cognition are overstated: they depend on a large inferential leap from one type of data about a particular type of causal cognition to a wide-ranging conclusion about the pervasiveness of normative considerations in causal cognition more generally. Proponents of these strong claims about causal cognition may respond that almost all empirical investigations involve some manner of ampliative

[1] "Normative" here refers broadly to many different types of norms, most notably moral or statistical norms, but also social norms. For ease of exposition, we will throughout subsume these various types of norms under the heading of "normative considerations."

[2] Roughly, some considerations are *constitutive* of a particular psychological process if and only if they are required inputs for this process, and if as a consequence they are always involved in it.

inference, and so it is insufficient for us to merely note that such inferences are occurring here as well. We thus describe some new empirical results showing that some types of causal cognition are *not* influenced by moral considerations.

2 An inference too far

The general thesis at issue—normative considerations influence causal cognition[3] and are perhaps even constitutive of various cognitive processes—can be understood in different ways, depending on one's positions vis-à-vis three issues. The first issue has received the most attention to date: namely, the nature and extent of the influence of normative considerations. There are many possibilities for the nature and degree of influence, but the most important split is arguably between those who hold that normative considerations are constitutive of the cognitive processes involved in causal judgments, and those who hold that they are *not* so constitutive. The first position has been defended by Hitchcock and Knobe (2009), who argue that causal judgments depend on counterfactual judgments, which, in turn, depend on normative judgments (see the quote in the introduction). On their view, use of the folk concept of causation in causal judgments depends on judgments about what is normal, and so normative judgments and reasoning are a constitutive part of our causal understanding of the world. There simply are no causal judgments—of any sort—that are free from normative influences. By contrast, for Alicke et al. (2011), causal judgments about human actions and the outcomes produced by those actions can be strongly influenced by the desire to blame or praise, but this influence occurs only after an initial causal judgment has been made (see also Alicke 1992, 2000). Normative considerations are thus not constitutive of the cognitive processes involved in causal judgments on this model since the initial, non-normative causal judgment is only modified in contexts involving blame or praise; other contexts should not exhibit any influence of normative considerations on causal judgment. The issue of the nature and extent of the influence of normative considerations on causal judgments is clearly philosophically significant, particularly for theories of causation; it bears, for example, on whether folk causal judgments can be used to support theories of causation (e.g., Machery 2011). Unsurprisingly, it has received significant attention among philosophers (see, e.g., many of the commentaries on Knobe 2010). The other two dimensions of the general thesis at hand have been less noticed.

The second issue concerns the nature of the normative considerations that are meant to influence causal judgments. We follow standard practice and distinguish here between *statistical* and *prescriptive* norms. The former provide the statistics of the environment, and so provide no "ought." For example, it is a statistical norm that more men than women are philosophers of science, even though most (perhaps all) would argue that there is no reason that there ought to be a gender imbalance.

[3] We will focus on everyday causal cognition in this article. It is unclear (even doubtful, in our opinion) whether proponents of the focal thesis intend it to apply to various types of causal inference and reasoning in the sciences, and so we leave that issue aside.

By contrast, prescriptive norms carry normative force, which itself can derive from many sources (e.g., one's social community or some other source). In addition to statistical and prescriptive norms, Hitchcock and Knobe (2009) argue that *norms of proper functioning*—norms about the operation of physical and biological systems organized to produce some outcome—are also relevant to causal judgments (see also McGrath 2005). There are thus many different possibilities for the set of normative considerations that can influence causal cognition, ranging from an exclusive focus on prescriptive norms (e.g., Alicke 1992; Sytsma et al. 2012) to a mix of all of these types of norms (e.g., Hitchcock and Knobe 2009). Variation in the scope of normative considerations thus leads to variation in the strength of the target claim.

The third issue has been almost entirely neglected: namely, the scope of "causal reasoning," "causal cognition," or "causal judgment." (All of these expressions are used in the philosophical literature, often interchangeably.) Causal cognition does not constitute a single cognitive activity, but rather includes several different cognitive processes and representations. A first distinction is between *causal learning* and *causal reasoning*, where the former is (roughly) the acquisition of causal knowledge and the latter is the use of that knowledge to generate predictions for novel situations, develop explanations for unexplained phenomena, or select one or another factor as *the* cause of some event (e.g., Ahn and Bailenson 1996; Hart and Honoré 1985).[4] There are at least three different processes through which people learn causal information. The first is *instruction*—written or oral—about a causal system; one might, for example, simply be told which light switch causes the lights to turn on. This type of learning has been surprisingly little studied in the context of causal knowledge (versus, say, learning scientific concepts), though there are results suggesting that it is not simply the straightforward acquisition of relevant facts (e.g., Taylor and Chi 2006). Second, *causal perception* is the relatively direct perception of causal relations between events in the world, such as seeing the cue ball cause the eight ball to move (Heider and Simmel 1944; Michotte 1946/1963; Scholl and Tremoulet 2000). Finally, *causal inference* refers to the relatively indirect learning of causal structure from statistical and other cues (e.g., Cheng 1997; Cheng and Novick 1992; Gopnik et al. 2004). Causal inference almost always involves learning a causal structure from a set of cases, whether presented serially or simultaneously. There are both behavioral (e.g., Schlottmann and Shanks 1992) and neuroscientific (e.g., Roser et al. 2005) results indicating that causal perception and causal inference are distinct cognitive processes (Danks 2009). Thus, any claim about the role of normative considerations in causal cognition must be clear about which cognitive processes are supposed to be affected.

These three issues give rise to an enormous space of possible claims about the connections between normative considerations and causal cognition. We suggest that the most interesting part of this space contains what we will call the "Ubiquity Thesis": Normative considerations (broadly construed) influence causal cognition

[4] This is not a clean division since one can learn new causal information by reasoning about existing knowledge (e.g., the so-called "self-explanation effect"; Chi et al. 1994; Van Lehn et al. 1992).

(broadly construed) and are perhaps even constitutive of various cognitive processes involved in aspects of causal cognition. This part of the logical space essentially consists of those hypotheses that assert that normative considerations are always or almost always relevant for all (or nearly all) aspects of causal cognition. All of the quotes with which we began this paper fit into this region of the logical space. Moreover, this region corresponds to the most substantive and surprising claims: it is not simply the minimal prediction that some normative considerations might sometimes matter for some aspects of causal cognition; rather, it is the sweeping claim that normative considerations are always, or almost always, relevant for essentially all causal cognition—a claim that challenges the arguably standard view in both philosophy and psychology that causal cognition is relatively free of normative influences.[5]

An influential argument among the defenders of the Ubiquity Thesis is that the empirical data compel it. We contend that there is no such compulsion. Essentially all of the experiments have had the same basic structure: participants are *told* about a situation (perhaps including some causal information), and then *asked* what caused some event. Participants complete the task by rating or ranking various causal candidates, by listing what they take to be the causes of the event, or some similar measure. The Ubiquity Thesis is then defended on the grounds that participants' normative evaluations (which are either measured during the task or simply hypothesized) influence their judgments about what event(s) in the situation caused some other event(s). For a representative example of the kinds of studies conducted in support of the Ubiquity Thesis, participants in Knobe and Fraser (2008) were given the following vignette:

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the 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 emailed them reminders that only administrative assistants 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.

After reading the vignette, participants were directly asked which of the following two sentences they agreed most with: "Professor Smith caused the problem" and "The administrative assistant caused the problem." Knobe and Fraser found that people were much more willing to say that the faculty member caused the problem, and concluded (p. 443) that "[t]he results therefore suggest that moral judgments actually do play a direct role in the process by which causal judgments are generated."

[5] For instance, McGrath (2005, p. 125) notes that "causation is commonly held to be a paradigmatic example of a natural and so entirely non-normative relation."

The empirical evidence offered in support of the Ubiquity Thesis is largely restricted to data concerning people's explicit, verbal judgments about specific, linguistically described cases. The existing body of evidence may well show that certain kinds of causal cognition—specifically, a highly language-driven kind of causal reasoning—are influenced by normative considerations, but it falls short of showing that causal cognition in general is influenced by normative considerations. A proponent of the Ubiquity Thesis could reply that it is natural to infer inductively from this body of evidence that normative considerations influence causal reasoning in general since there is no immediate reason to think that other aspects of causal reasoning differ from those about which we have direct evidence. However, as we argue in the next section, these experiments involve so-called "learning from description," and the features of the cognitive processes engaged by this kind of task are unlikely to generalize to all kinds of causal cognition. That is, we actually do have principled reasons to be concerned about the inductive move from these experiments to all causal cognition, and so the extant experiments do not give us sufficient positive reasons to accept the Ubiquity Thesis.

Defenders of the Ubiquity Thesis have additionally offered theoretical arguments for this thesis. For example, Hitchcock and Knobe (2009) and Knobe (2009) argue that causal judgments are partly based on counterfactual judgments, which are partly based on judgments about normality. That is, counterfactual judging is always a part of all causal judging. We must be careful with this argument, however, since the link between counterfactual and causal judgment is not as strong as many philosophers think. Despite the close connections between counterfactuals and causation in philosophical accounts, there are substantial psychological differences between causal learning and reasoning, and counterfactual judgment (Mandel 2003). In particular, there is substantial evidence that causal learning and reasoning do not always involve counterfactual judging (Danks 2009). For example, experimental participants will (spontaneously) attribute significant causal efficacy to factors that do not enter into any (spontaneously generated) counterfactual statements, contrary to what we should expect if counterfactual judging were a constitutive part of causal judgment. Of course, counterfactual judgments surely influence *some* of our causal reasoning, and those judgments perhaps depend *in part* on judgments of normality. But this is much less than is required for the Ubiquity Thesis to hold.

3 Positive evidence for a negative thesis

We have so far provided only a negative argument: we have shown that the theoretical and empirical reasons offered in defense of the Ubiquity Thesis do not establish that normative considerations influence all causal cognition. We have not yet shown, however, that the Ubiquity Thesis is actually false; its defenders have made unwarranted inferential leaps, but lack of warrant does not imply falsity. We thus turn in this section to providing positive reasons, both theoretical and empirical, to think that the Ubiquity Thesis is false. We will not dispute the empirical evidence that normative considerations can sometimes influence certain narrow aspects of

causal cognition. We concede that there seem to be particular instances of causal instruction and reasoning that are influenced by normative considerations. Our contention is instead that these effects are highly limited in scope.

From a theoretical point of view, causal perception appears to be a quite different cognitive process than other types of causal cognition. It is largely centered in the visual cortex (Fugelsang et al. 2005), automatic, and not subject to top-down control (Blakemore et al. 2001). These aspects have led some to suggest that causal perception might even be a Fodorian module (Scholl and Tremoulet 2000; though see Schlottmann 2000). Thus, to the extent that normative considerations can play a role only for higher-level, more "cognitive" processes, we should not expect them to have a significant effect on causal perception.[6]

The defender of the Ubiquity Thesis could respond to these observations in at least two different ways. First, one could observe that normative cognition is itself modular in some respects, and so could influence lower-level processes such as causal perception; it need not play a role only in higher cognitive processes. In support of the modularity of normative cognition, one could note that normative judgments are often cognitively impenetrable—one has no access to the processes outputting these judgments (Haidt 2001; Cushman et al. 2006)—and that normative cognition develops early (Cummins 1996; Rakoczy et al. 2008). In response, we note first that the modularity of normative cognition would at best show that it *could* influence causal perception, not that it *does*. Second, even if normative cognition is modular in some respects, it still must operate incredibly fast (on the order of 100 ms, if not faster) in order to have an impact on causal perception. Unfortunately, there are no experiments directly measuring the time course of normative evaluations, only ones studying rapid decisions in normatively charged contexts (e.g., Suter and Hertwig 2011). Nonetheless, although modularity typically implies some increase in processing speed, it is prima facie implausible that normative evaluations would be faster than perceptual processes, which is, at the very least, what would be required for normative evaluations to influence causal perception.

Second, the defender could argue that perceptual processes in general can be penetrated by higher cognition, as proponents of cognitive penetration have argued (e.g., Stokes 2012; Macpherson 2012), and that causal perception itself may thus be cognitively penetrable. First, this response depends on the controversial claim that early perceptual processes are cognitive penetrable, and there are good reasons to resist this general claim (see Machery ms). Second, this response would again establish, at best, that normative cognition *could* influence causal perception, not that it *does*. Third, even if causal perception is cognitively penetrable, the pervasive nature of causal illusions (e.g., Michotte 1946/1963; Scholl and Tremoulet 2000)

[6] Normative evaluations could possibly have an impact on low-level processing by changing the individual's affective state, which then might influence causal perception. Causal perception is incredibly rapid, however, and it seems unlikely that the other processing could occur fast enough to matter. Moreover, even if this theoretically possible "back door path" existed, it would be restricted to normative evaluations that have an emotional impact (e.g., not most evaluations of statistical typicality).

suggest that this penetrability must be quite limited: much of causal perception is clearly not penetrated by our beliefs.

Even if these responses to the causal perception-based objection to the Ubiquity Thesis do not work, it is unclear whether the objection should really bother a proponent of the Ubiquity Thesis. She could quite naturally concede that causal perception is not influenced by normative considerations, while insisting, with some justice, that the Ubiquity Thesis was meant to apply to higher cognition in general, viz. to processes resulting in causal judgments rather than causal percepts, including causal inference and causal reasoning. Causal inference does not have the same type of automaticity as causal perception, but rather takes place over an extended period of time, involves relatively high-level cognitive and neural processing (Satpute et al. 2005), and is subject to significant top-down influences of domain knowledge (Schulz et al. 2007), temporal information (Buehner and May 2002), and knowledge of underlying mechanisms (Schlottmann 1999). The door is thus clearly open for normative considerations to influence causal inference. At the same time, we should not jump to the conclusion that such effects actually do occur.

A helpful lesson comes from the research on contingency and probability learning—how do people determine the probability of an event?—and on choice under uncertainty—how do people choose between courses of action whose outcomes are only probable? In both fields, a critical distinction is drawn between learning from description and learning from experience: the former refers to simply being told various probabilities, while the latter refers to learning the relevant probabilities through experiences and interactions. A large body of research shows that these two forms of learning are very different (e.g., Barbey and Sloman 2007; Erev et al. 2010; Gigerenzer and Hoffrage 1995; Hau et al. 2008, 2010; Hertwig and Erev 2009). In choices under uncertainty, for example, very unlikely outcomes are underweighted when their probability is learned from experience, but overweighted when their probability is learned from description (Hau et al. 2008). In contingency learning, people make basic judgment and estimation errors after learning from description that they do not make after learning from experience (Barbey and Sloman 2007). A common explanation is that people who learn from description typically engage in explicit, high-level reasoning that is slow, error-prone, and subject to outside influences. In contrast, those who learn from experience use other reasoning processes—sometimes described as heuristics (e.g., Hertwig and Erev 2009) or system-1 processes (Evans 2008)—that are not subject to these influences. Essentially all of the experiments supporting the Ubiquity Thesis have focused on learning from description: participants are explicitly told the relevant causal structure. By contrast, causal inference is almost always understood as a process of learning from experience: people learn the causal structure and the strengths of the causal relations from a set or sequence of cases rather than direct instruction. Perhaps causal inference is, like learning from experience in contingency learning and choice under uncertainty, resistant to outside influences, including normative considerations.

Given these reasons to doubt that normative considerations influence causal inference, we should decide the issue empirically. We conducted two separate experiments that were structurally identical (see "Appendix 1" for the full

experimental materials and results). In each, participants were presented with two different scenarios; for each, they were told a general cover story, provided with a sequence of observational data about a potential cause and a target effect, and then asked to judge causal strengths (and to answer two additional questions in one scenario). The *control scenario* used a neutral cover story—specifically, about the influence of a plant on rashes—that had previously been used in causal inference experiments (Danks and Schwartz 2005, 2006). The *moralized scenario* introduced a significant normative component, as the cover story focused on an individual who was engaged in morally reprehensible activities (specifically, trying to destroy a cure for cancer). Each participant saw both the control and moralized scenarios (in random order); for each scenario, the sequence of cases led participants to infer a generative (the target cause produces the target effect), preventive (the target cause prevents the target effect from occurring), or no (the target cause is irrelevant to target effect) causal relation.[7] After seeing the full sequence, participants were asked to rate (on a standard $[-100, +100]$ scale) the causal strength of the potential cause. For the moralized scenario, participants were additionally asked to assess the blameworthiness of the morally bad actor (i.e., a "blame judgment"), and whether he knew the likely effect of his actions (i.e., a "knowledge judgment").[8]

The Ubiquity Thesis predicts that normative considerations should affect causal inference similarly to their significant influence on explicit, verbal causal judgments in standard learning from description studies (e.g., Alicke 1992). In particular, the Ubiquity Thesis predicts that there should be a difference between the rating of causal strength for a sequence in the control scenario and the rating for the same sequence in the moralized scenario. However, no such effect was found across the two experiments: the sequence-type (generative vs. preventive vs. non-causal) made a highly significant difference in the causal strength ratings, but there was no difference in ratings between the control and moralized scenarios.[9] That is, both scenarios yielded the same pattern of causal ratings that one finds in all standard causal inference experiments. Participants responded appropriately to variations in the statistics of the case sequences, but were apparently uninfluenced by their moral evaluations of the agent in the story.

One possible concern is that the cover story in the moralized condition might not have induced a sufficient blaming reaction. There is no evidence of this in the data,

[7] We used standard case sequences that have been used in other causal inference experiments. For each sequence, cases were presented in the same pseudo-random order for each participant. $Freq(C) = .5$ in all sequences, and the conditional frequencies of E were: *Generative*: $Fr(E \mid C) = .75$ and $Fr(E \mid \neg C) = .25$; *Preventive*: $Fr(E \mid C) = .25$ and $Fr(E \mid \neg C) = .75$; and *Non-causal*: $Fr(E \mid C) = Fr(E \mid \neg C) = .5$.

[8] In order to be able to sensibly present all three sequence-types in the moralized scenario, we used a cover story in which the morally bad actor did not actually know the causal efficacy (or even direction) of his actions, but was so desperate that he tried it anyway.

[9] We performed targeted ANOVAs that included only sequence-type, condition, and their interaction; for both experiments, sequence-type was the only significant predictor (Experiment #1: $F = 49.3689$, $p < 10^{-15}$; Experiment #2: $F = 26.5692$, $p < 10^{-9}$). Tukey HSD post hoc tests showed that all three sequence-types were significantly different from one another, which is relatively clear simply from looking at the mean ratings (Exp. #1: $Gen = 45.2$, $Non\text{-}causal = 7.5$, $Prev = -26.2$; Exp. #2: $Gen = 49.55$, $Non\text{-}causal = 4.35$, $Prev = -19.84$).

however.[10] A more subtle concern is that perhaps the influence of the moral evaluation depends on the intensity of the blaming reaction, and that any effects are being "washed out" because only *some* people exhibited a sufficiently strong reaction. Similarly, perhaps beliefs about the agent's knowledge (in particular, whether he knew the likely effect of his actions) played a role in people's causal inferences and thereby masked the effects of the moral evaluation. If either concern were correct, then participants' blame or knowledge judgments should predict their causal strength rating. But neither did.[11] Although it is notoriously difficult to establish the null result that normative considerations do *not* have an effect, there simply does not seem to be any evidence that normative considerations influenced causal inference in these experiments.[12] These experimental results converge with the theoretical considerations we presented above: even if some kinds of causal cognition are influenced by normative considerations, much of causal cognition likely is not, and the Ubiquity Thesis is probably mistaken.

4 Objections and replies

We now consider several objections, along with our replies, to the case against the Ubiquity Thesis that we have just presented. First, one could object that a within-subjects design is not the best design for testing the Ubiquity Thesis: the very act of reading both a moral and a non-moral scenario may have prompted participants (possibly unconsciously) to override the natural tendency to let normative considerations influence their causal judgments. In response, we note first that some within-subjects studies still show an influence of normative considerations on people's judgments (e.g., Pinillos et al. 2011 study of the Knobe effect), so there is no *a priori* reason to think that a within-subjects design cannot work. Second, participants knew only that they would see two scenarios, not that one would be moral and one non-moral. Thus, there should not be any self-correction (even unconsciously) in the first rating provided by each participant (i.e., the causal strength rating for whichever scenario was seen first), and so if the Ubiquity Thesis is correct, we should see an effect of scenario-type when we compare these ratings. Again, however, sequence-type (i.e., the observed evidence) was the only significant factor.[13]

[10] The mean blame ratings were 7.87 (Exp. #1) and 5.88 (Exp. #2) on a 1–9 scale, where higher numbers indicate more blameworthy. One complication emerged for Experiment #2: blame ratings differed between the generative (mean = 4.56) and preventive (mean = 7.06) conditions (Tukey HSD post hoc test yields $p < .01$). This suggests that perhaps both outcome and intentions can matter for blame.

[11] We performed ANOVAs for the moralized condition with sequence-type, blame judgment, knowledge judgment, and all interactions. The only significant predictor for either experiment was sequence-type (Exp. #1: $F = 14.6268$, $p < 10^{-5}$; Exp. #2: $F = 6.3084$, $p < .005$). Note that there was substantial variation in the knowledge ratings (Exp. #1: $\sigma = 2.38$; Exp. #2: $\sigma = 2.10$).

[12] Power calculations are provided in the next section.

[13] In ANOVAs for rating given sequence-type and condition, sequence-type was the only significant predictor (Exp. #1: $F = 42.6323$, $p < 10^{-7}$; Exp. #2: $F = 32.5542$, $p < 10^{-6}$).

Second, one could question whether the moralized scenario in Experiment 2 really elicited a negative moral judgment since the mean blame response—5.88 on a 9-point scale with higher numbers representing more blame—is close to the midpoint. This blame rating is nonetheless significantly greater than the mid-point (one-sample t test comparing against $\mu = 5$: $t = 2.4522$, $p < .01$); the mean response does not seem to be neutrality. Moreover, there is no question that there was a significant blaming response in Experiment 1 (mean was 7.87), and the two studies yield qualitatively identical results. It is unlikely that an insufficient blaming response can explain away both sets of results.

Third, one could object that inferences to a conclusion on the basis of null results (i.e., lack of effect) are either always invalid (following Fisher 1935), or at least extremely difficult to draw validly. We will bracket the first objection since this is not the place to defend the legitimacy of inferences from negative results and since one of us has done it at length elsewhere (Machery 2012). To address the second objection, we note first that our two experiments had enough power to detect the effect of sequence-type on participants' causal judgments, so any effect of normative considerations must be much smaller than the effect of sequence-type. Second, more formally, we can compute the power of our experiments, which gives the probability of correctly rejecting the null hypothesis of no effect when it is actually false—P(reject null hypothesis | null is false). The power of an experiment depends on the size of the effect if there really is one (i.e., what happens if the null hypothesis is false). No previous studies have attempted to measure the effect size (Cohen's f) of moral considerations in an experiment like this one, so we do not have *a priori* estimates to use for our power calculations. As an example, however, the effect size of the sequence-type in Experiment 2 is $f = .73$. Experiment #1 had 63 participants and Experiment #2 had 48 participants (see "Appendix 1"). For these sample sizes, power for a moderate ($f = .4$) effect of scenario-type is .99 (Exp. #1) and .97 (Exp. #2); for a small ($f = .25$) effect, power is .80 (Exp. #1) and .68 (Exp. #2).

Fourth, and perhaps most importantly, one could argue that the Ubiquity Thesis does not assert that *all* moral judgments affect causal inference and causal reasoning. That is, it is consistent with the Ubiquity Thesis that some moral judgments do not affect causal inference and causal reasoning. And the worry is that our two experiments might have elicited normative reactions of exactly this type: perhaps the morally blameworthy actions in the moralized scenario do not influence *any* causal cognition, whether causal inference (as in our experiments) or causal reasoning after learning from description (as in most previous experiments on this issue).

In response, we first note that our moralized scenario is quite similar in structure and content to other vignettes used to show the influence of moral judgments on causal reasoning (see "Appendix 1, 2" for full materials). There is no obvious reason to think that the normative reactions elicited in these experiments are of a different type than those elicited in other experiments.

Nonetheless, simply pointing to vignette similarity is a relatively weak response, as the question of which vignette features actually elicit normative reactions of particular types is itself an open research problem. We thus ran an additional

experiment that uses these two scenarios in a learning by description task; in other words, we used our scenarios in an experiment that is as similar as possible to the type of studies that have been used to support the Ubiquity Thesis. If our vignettes failed to elicit the "right" normative reaction (viz. the kind of normative reaction that influences causal judgment), then we should again find no differences in causal judgment between the two scenarios. But if we instead find the "standard" between-scenario judgment differences, then it seems clear that our moralized scenario must be producing the "right" type of normative reaction, and so the lack of between-scenario differences would have to be explained in some other way.

We presented each participant with one of the scenarios used in Experiment 1 (see "Appendix 2" for full materials and analyses). We could not, however, show them a series of cases, as this experiment was intended to be purely verbal. We instead simply told participants that either "most," "few," or "about half" of the plants were dead or people had rashes, respectively. This resulted in a two (Scenario: Moralized, Control) by three (Outcome: Most, Few, Half) between-subjects design. For all conditions, after the participant read the vignette they were asked to rate whether the causal candidate caused the effect on a scale ranging from -10 to 10 with -10 anchored with "always prevents," 0 anchored with "irrelevant," and 10 anchored with "always produces." In addition, participants who read the moralized scenario were asked to make blame and knowledge ratings, using the same 9-point scale as in our previous experiments.

The results exhibit a clear influence of normative considerations.[14] Specifically, for each of the three Outcome possibilities, participants gave significantly higher causal ratings in the moralized scenario than in the control scenario. These results suggest that the moral considerations present in our scenarios do influence people's causal judgments when learning from *description*. Thus, our previous null results cannot be explained away by appealing to the idea that according to the Ubiquity Thesis not all normative considerations influence judgment and to the hypothesis that the moral consideration we used are not of the "right" kind. Rather, it seems most plausible that normative considerations do not actually influence causal inference (i.e., learning from experience). Thus, we have positive evidence that the Ubiquity Thesis is, in fact, an inference too far.

5 Conclusion

Recently, philosophers and psychologists have claimed that normative considerations, both moral and non-moral, play a fundamental role in causal cognition. There are a variety of ways to clarify this vague claim, and we focused on its strongest, but also most frequently endorsed, version—the Ubiquity Thesis. This thesis asserts that normative considerations, broadly understood, fundamentally influence causal cognition in general, perhaps because they are constitutive of the

[14] In an ANOVA for Rating given Scenario and Outcome, there were main effects of both Scenario ($F = 26.615$, $p = .000$) and Outcome ($F = 43.252$, $p = .000$). A significant main effect of Scenario was also found in each individual Outcome condition (see "Appendix 2").

cognitive processes. However, the Ubiquity Thesis is probably false, and the connections between normative considerations and causal cognition are more complex than recent discussions assume. Both the empirical evidence and the various theoretical arguments for the Ubiquity Thesis are lacking. This thesis is supported only by a narrow type of data from causal instruction experiments that are likely to tap into a peculiar kind of causal cognition. Moreover, our two causal inference experiments strongly suggest that normative considerations do not influence causal inference. It may thus well be that the influence of normative considerations is limited to a narrow, linguistically-mediated form of causal cognition.

Appendix 1

Experiment 1

Participants

Sixty-three undergraduates at Carnegie Mellon University participated in return for $5. The experiment took approximately 20 min to complete.

Materials and methods

Participants were first provided with a general overview, followed by the Moral and Control scenarios (randomized presentation order). For each scenario, participants were first provided with a cover story, and then observed one of three possible 48-case sequences—either a generative, neutral, or preventive sequence—where the abstract structure was the same regardless of scenario.[15] To eliminate possible familiarity effects, participants saw different (abstract) sequence-types in the different scenarios, and so there were six different possible conditions (e.g., generative in the Moral condition and preventive in the Neutral condition). The frequency distributions for the three 48-case sequences were:

- *Generative*: $P(C) = .5$; $P(E \mid C) = .75$; $P(E \mid \neg C) = .25$
- *Neutral*: $P(C) = .5$; $P(E \mid C) = P(E \mid \neg C) = .5$
- *Preventive*: $P(C) = .5$; $P(E \mid C) = .25$; $P(E \mid \neg C) = .75$

For each sequence, participants observed the cases one at a time, and clicked a button to move to the next case. Each case was described both in text (e.g., "Betrafindalis copernicia: Alive") and with corresponding images. After observing the full sequence, participants were asked to rate "to what extent [the potential

[15] The 48-case sequences used were:
- *Generative*: 010203112010310110023110310010210011302110021130
- *Neutral*: 013220313200213103221013201033121302032103212130
- *Preventive*: 232021330232132332201332132232032233120332203312 where '0' denotes a *C&E* case, '1' is $\neg C\&\neg E$, '2' is $C\&\neg E$, and '3' is $\neg C\&E$.

cause] causes [the target effect].” Ratings were collected using a slider that ranged from −100 (“Always prevents”) to +100 (“Always produces”) with an anchor at 0 (“No effect”). The slider moved in increments of 5, so actually corresponded to a 21-point scale. The slider began at 0, but had to be moved before the rating could be submitted (i.e., participants could not simply click through without moving the slider).

The global introduction was:

> You are about to be presented with two stories. One story is about a man, Smith, and plants that he is growing in his greenhouse. In the story, a liquid has been applied to the plants but the characters in the story do not know what liquid has been applied to the plants. The liquid may either be a fertilizer, poison or water, and so might lead the plants to flower, die, or it has no effect at all. Your job will be to figure out what liquid was applied to the plants.
>
> You must remember that the relationship between the plant dying or flowering and exposure to the liquid could be quite complicated (if there is any relationship at all!). As an example, there are many plants that are very sensitive to fertilizers, and flower very easily if exposed to them. But, some plants that are very sensitive to fertilizers still might not flower when exposed to them. Likewise, there are many plants that are very resilient and thus do not have serious reactions when exposed to a poison. But plants that are resilient may still have a serious reaction when exposed to a poison.
>
> The other story that you will see is about Johnson, a doctor who has traveled to an island to study the outbreak of a skin disease among a particular group of villagers. Villagers have come into contact with various plants on the island and some have contracted rashes. Your job will be to figure out whether exposure to a certain plant causes the skin disease, makes people healthy, or has no effect at all.
>
> You must remember that the relationship between the rashes (a symptom of the disease) and exposure to the plant could be quite complicated (if there is any relationship at all!). And this skin disease is like many other diseases: different villagers might have different levels of immunity or resistance, and there are likely many different causes of the disease. As an example, there are many people who respond readily to vitamins, and very easily become healthy if they take them. But, in some cases people who normally respond to vitamins may still not become healthy when exposed to them. Likewise, there are many people who are allergic to peanuts, and break out in serious reactions if exposed to them. But, in some cases, people who are allergic to peanuts might not have a serious reaction when exposed to them.
>
> For both stories, you will be presented with a series of individual cases. For each case, you will be shown whether the factor (liquid or plant) was present or absent, and what happened to the flower or person. The factor’s absence

> will be indicated by a red X over the picture of the factor. These cases will help you figure out whether or not exposure to a particular liquid causes the plant to die, causes it to grow, or has no effect at all; and whether exposure to a particular plant causes villagers to contract a skin disease, causes them to be healthy, or has no effect at all. After viewing all of the pictures, you will be asked to evaluate the causal connection between these factors.

The cover story for the Control scenario was:

> Johnson is a doctor traveling to the South Pacific Islands to research the rare skin disease Anthrapora that has been reported on various islands. In particular, she is studying the possible effect of native plants have on the contraction of these diseases. On the island of Tongatapu, Johnson is studying the impact (if any) of *Solanaceae delisa* on the skin disease Anthrapora. The plant may lead to the skin disease, it may cure the disease, or it may have no real effect at all. It is your job to figure this out.
>
> Johnson interviewed various villagers; some have the local disease, and some do not. She can diagnose villagers as suffering from the skin disease by finding the characteristic rashes. Unfortunately, because of language barriers, the only other information she can get from the villagers is whether or not they have come in contact with the local plant, *Solanaceae delisa*.
>
> There are thus four different observations Johnson might make: the villager was exposed to the plant and suffers from the disease; the villager was exposed to the plant and is healthy; the villager was not exposed to the plant and suffers from the disease; the villager was not exposed to the plant and is healthy.
>
> You will now see the information – both plant contact and disease status – that Johnson collected for several villagers. After seeing all of the individuals, you will be asked to evaluate the causal connection between the plant and the skin disease on a scale from −100 to +100. Respond with −100 if you think that exposure to the plant (*Solanaceae delisa*) always prevents the skin disease (Anthrapora). Respond with +100 if you think that exposure to the plant always produces the rashes. And respond 0 if you think the plant is irrelevant for whether the person suffers from the disease or is healthy. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

The cover story for the Moralized scenario was:

> Smith is an elderly man who has devoted his life to cancer research. He has been involved in the development of various treatments, which have helped to save the lives of thousands of people. Recently, Smith traveled deep into the Amazon in order to recover a nearly extinct species of orchids called Betrafindalis copernicia.
>
> Betrafindalis copernicia contains a highly concentrated form of the chemical dispofignila. Smith has been experimenting with a synthetic form of dispofignila and has found that it slows down the growth of cancer. While

synthetic doses slow the growth down somewhat, only the strongest form of dispofignila, which cannot be synthetically produced and is only found in the orchid Betrafindalis copernicia, slows down the growth substantially, almost to the point of stopping growth altogether. As a matter of fact, Smith was actually able to experiment with one Betrafindalis copernicia. He found that the dispofignila found in the Betrafindalis copernicia was in fact more potent than the synthetic form of dispofignila and that it lead to a significantly greater decrease in cancer growth than the synthetic form of dispofignila.

Betrafindalis copernicias are very rare. They are only found in a remote part of the Amazon, and, because of global warming, only a few dozen plants survive even there. Smith is sure that if he can preserve these plants, then he can develop a cure for cancer. Thus, Smith traveled to the Amazon, returned home safely with all of the orchids, and placed them in his greenhouse.

Smith's neighbor, Jones, hates Smith. Jones has always despised Smith for no good reason. Jones knows that Smith has recently returned with the only remaining orchids in the world and he wants to kill all of the plants and destroy Smith's hopes for finding a cure for cancer.

Jones wants to kill the orchids, but he doesn't want to get caught. He knows that he could simply uproot the plants and kill them, but in order to do that he would have to get into the greenhouse, which is secured by an alarm. However, there is one way that Jones can kill the plants without leaving any evidence. There are several hoses that run a steady flow of water into the greenhouse. Jones knows that if he can inject poison into the hoses, it will kill all of the plants, and no evidence will be left behind.

One night, Jones breaks into an old farmer's shed and finds several bottles with the label "poison"; Jones grabs one of the bottles and discretely leaves the shed. Unbeknownst to Jones, however, the farmer reuses his bottles: some of the ones labeled "poison" have fertilizers, others simply have water, and of course, some actually do have poison. (The farmer has a system for knowing what each bottle contains, though Jones obviously does not know this system.)

A few days later, Jones fills a syringe with the liquid from the stolen bottle. He then goes to Smith's greenhouse, pushes the needle through one of the hoses, and injects the liquid into that particular hose.

Your job is to find out whether the liquid that Jones injected into the hose was a poison, a fertilizer, or water (in which case it has no effect). There are thus four different observations you might make: the plant was exposed to the liquid and dies; the plant was exposed to the liquid and produces flowers; the plant was not exposed to the liquid and dies; the plant was not exposed to the liquid and produces flowers.

You will see the relevant information – both liquid contact and whether the plant died or flowered – for several of the plants in Smith's greenhouse. After seeing all of the plants, you will be asked to evaluate the causal connection

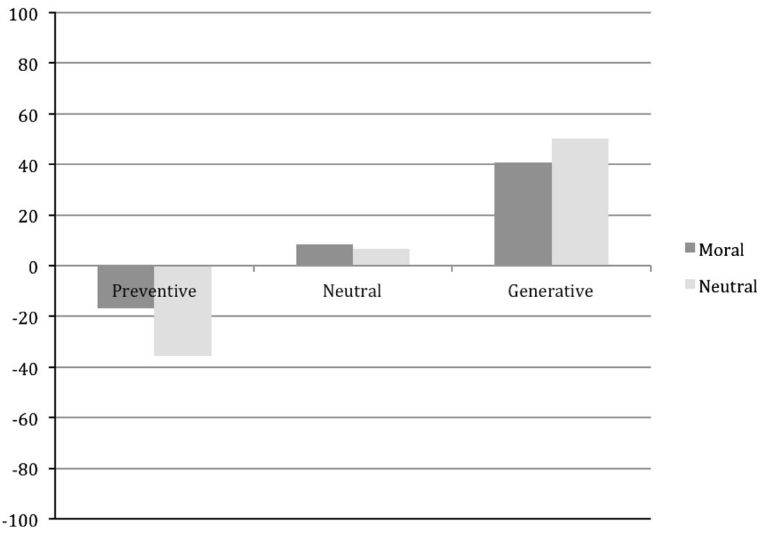


Fig. 1 Mean strength ratings in each condition

between the liquid and the plant flowering on a scale from −100 to +100. Respond with −100 if you think that exposure to the liquid always kills the plant. Respond with +100 if you think that exposure to the liquid always makes the plant flower. And respond 0 if you think the liquid is irrelevant for whether the plant dies or flowers. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

After only the Moralized scenario, participants were asked two further questions with responses on a 9-point scale: a Blame question ("How blameworthy do you think Jones is for attempting to kill Smith's plants?" from "Not at all blameworthy" to "Extremely blameworthy") and a Knowledge question ("Do you think that Jones knew what was in the bottle of liquid that he stole?" from "He did not know at all" to "He definitely knew").

Results and analysis

There were no order effects in the data, so we report results pooling different orders together. The mean ratings are shown in Fig. 1. ANOVA revealed that there was a main effect of sequence-type ($F = 49.3689$; $p = 2.267 \times 10^{-16}$), but no effect of situation ($F = .7742$; $p = .3807$) and no interaction effect ($F = 1.9480$; $p = .147$). Tukey HSD post hoc tests showed that all three sequence-types are significantly different from one another (G vs. P: $p = .0$; G vs. N: $p = 2.1 \times 10^{-6}$; N vs. P: $p = 2.34 \times 10^{-5}$).

Participant ratings in the moralized scenario are perhaps influenced by the participant's views about blame or knowledge. An ANOVA for Rating with inputs of sequence-type, blame and knowledge judgments, and all interactions showed that

sequence-type was the only significant predictor ($F = 14.6268$; $p = 9.532 \times 10^{-6}$); all other p-values were at least .05. Tukey HSD post hoc tests showed that the Generative sequence-type led to significantly different ratings than the Neutral and Preventive sequence-types in the restricted domain of only the Moral condition, and the latter two sequence-types were close-to-significantly different (G vs. P: $p = .0000069$; G vs. N: $p = .014$; N vs. P: $p = .0743$).

An ANOVA of Blame judgments using sequence-type, knowledge judgments, strength ratings, and all interactions found no significant predictors (all p-values $> .12$). An ANOVA of Knowledge judgments using sequence-type, blame judgments, strength ratings, and all interactions found a significant effect of only the Sequence-type $\times$ Rating interaction ($F = 4.5785$; $p = .01483$); all other factors were not significant (all p-values $> .10$). Specifically, knowledge judgments were negatively correlated with rating for preventive sequences, positively correlated for neutral sequences, and essentially uncorrelated for generative sequences. We have no clear explanation for this small effect.

Experiment 2

Participants

Forty-eight undergraduates at Carnegie Mellon University participated in return for $5. The experiment took approximately 20 min to complete.

Materials and methods

The basic method of Experiment 2 was identical to Experiment 1, and differed only in the cover stories that were used. The global introduction was:

> You are about to be presented with two stories. One story is about a man, Jones, and plants that his competitor is growing. In the story, Jones prunes some of the plants, which may help them survive, may kill them, or may have no effect at all.
>
> You must remember that the relationship between the plant living or dying, and the plant being pruned could be quite complicated (if there is any relationship at all!). As an example, there are many plants that require pruning to survive and thrive, while other plants are very sensitive and so die when pruned. And there are many plants that are very resilient and thus do not normally have any significant reaction when pruned (but might in any particular case).
>
> The other story that you will see is about Johnson, a doctor who has traveled to an island to study the outbreak of a skin disease among a particular group of villagers. Villagers have come into contact with various plants on the island and some have contracted rashes. Your job will be to figure out whether exposure to a certain plant causes the skin disease, makes people healthy, or has no effect at all.

> You must remember that the relationship between the rashes (a symptom of the disease) and exposure to the plant could be quite complicated (if there is any relationship at all!). And this skin disease is like many other diseases: different villagers might have different levels of immunity or resistance, and there are likely many different causes of the disease. As an example, there are many people who respond readily to vitamins, and very easily become healthy if they take them. But, in some cases people who normally respond to vitamins may still not become healthy when exposed to them. Likewise, there are many people who are allergic to peanuts, and break out in serious reactions if exposed to them. But, in some cases, people who are allergic to peanuts might not have a serious reaction when exposed to them.
>
> For both stories, you will be presented with a series of individual cases. For each case, you will be shown whether the factor (pruning or plant) was present or absent, and what happened to the plant or person. The factor's absence will be indicated by a red X over the picture of the factor. These cases will help you figure out whether or not pruning causes the plant to die, causes it to grow, or has no effect at all; and whether exposure to a particular plant causes villagers to contract a skin disease, causes them to be healthy, or has no effect at all. After viewing all of the pictures, you will be asked to evaluate the causal connection between these factors.

The cover story for the Control scenario was:

> Johnson is a doctor traveling to the South Pacific Islands to research the rare skin disease Anthrapora that has been reported on various islands. In particular, she is studying the possible effect of native plants have on the contraction of these diseases. On the island of Tongatapu, Johnson is studying the impact (if any) of *Solanaceae delisa* on the skin disease Anthrapora. The plant may lead to the skin disease, it may cure the disease, or it may have no real effect at all. It is your job to figure this out.
>
> Johnson interviewed various villagers; some have the local disease, and some do not. She can diagnose villagers as suffering from the skin disease by finding the characteristic rashes. Unfortunately, because of language barriers, the only other information she can get from the villagers is whether or not they have come in contact with the local plant, *Solanaceae delisa*.
>
> There are thus four different observations Johnson might make: the villager was exposed to the plant and suffers from the disease; the villager was exposed to the plant and is healthy; the villager was not exposed to the plant and suffers from the disease; the villager was not exposed to the plant and is healthy.
>
> You will now see the information – both plant contact and disease status – that Johnson collected for several villagers. After seeing all of the individuals, you will be asked to evaluate the causal connection between the plant and the skin disease on a scale from −100 to +100. Respond with −100 if you think that exposure to the plant (*Solanaceae delisa*) always prevents the skin disease

(Anthrapora). Respond with +100 if you think that exposure to the plant always produces the rashes. And respond 0 if you think the plant is irrelevant for whether the person suffers from the disease or is healthy. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

The cover story for the Moralized scenario was:

Miracle Works is a cancer research company that investigates the potential effects that various chemicals produced by exotic plants have on cancer cell growth. Recently, a chemical was isolated in a plant called Bertandis Capernicalia. This chemical has been extensively studied and has been shown to reverse the growth of cancer cells. However, the chemical is only produced after the plant has been fully mature for 2 weeks.

The plant was discovered by researchers after they crossed two different plants, Detra Nicalia and Berta Capernica. Researchers soon found out, however, that Bertandis Capernicalia was very hard to keep alive after it was fully matured. Researchers think that pruning might matter, but they do not know exactly how, since the two plants crossed to produce Bertandis Capernicalia—Detra Nicalia and Berta Capernica—respond differently to pruning. Detra Nicalia lives longer if darkish green leaf tips are trimmed off; Berta Capernica dies faster when it is pruned. Thus, pruning might help Bertandis Capernicalia, harm it, or turn out to be just irrelevant.

Jones is a head executive for CFF Treatment Inc., which is a large chemotherapy company. CFF Treatment Inc. makes billions of dollars a year through manufacturing various machines and chemicals that are used in chemotherapy. Jones and the other executives of CFF Treatment Inc. know that if Miracle Works can successfully grow Bertandis Capernicalia, then cancer patients may be able to be effectively treated and cured without ever having to go through chemotherapy.

Jones decides that something must be done in order to prevent Miracle Works from successfully growing Bertandis Capernicalia. Jones decides that Miracle Works must be sabotaged and knows that he must be clever so that he is not caught. He decides that the best way to sabotage Miracle Works (and not get caught) is to prune all of the plants. This way, it will look like one of the Miracle Works employees is responsible for the death of all the plants.

Importantly, Jones has no idea whether pruning will help or hurt the plants. He thinks that it will effectively kill all of the plants, but this is just a guess on his part. Recall that not even the scientists who work at Miracle Works are sure of the effects of pruning of Bertandis Capernicalia.

One night, Jones breaks into Miracle Works and starts to cut the tips off of all Bertandis Capernicalia leaves. A Miracle Works employee walks in and so he is interrupted and must sneak out before he is caught. Due to the interruption, he was only able to prune some of the plants.

> You will now see pictures of plants that were randomly chosen from Miracle Works. Some plants have had the tips cut while others have not. Some plants survived, and some did not. Your job will be to determine whether pruning harms, helps, or is irrelevant to Bertandis Capernicalia. After seeing all of the plants, you will be asked to evaluate the causal connection between pruning and the plant surviving on a scale from −100 to +100. Respond with −100 if you think that pruning always kills the plant. Respond with +100 if you think that pruning always makes the plant survive. And respond 0 if you think that pruning is irrelevant for whether the plant lives or dies. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

Suitably adjusted Blame and Knowledge questions were again asked after the causal strength rating was elicited in the Moral situation.

Results and analysis

There were no order effects in the data, so we report results after pooling together different orders. The mean ratings are shown in Fig. 2. ANOVA revealed that there was a main effect of sequence-type ($F = 26.5692$; $p = 8.548 \times 10^{-10}$), but no effect of situation ($F = .0109$; $p = .9172$) and no interaction effect ($F = 1.8921$; $p = .1567$). Tukey HSD post hoc tests showed that all three sequence-types are significantly different from one another (G vs. P: $p = .0$; G vs. N: $p = .0000372$; N vs. P: $p = .042$).

Participant ratings in the moralized situation are perhaps influenced by the participant's views about blame or knowledge. An ANOVA for Rating with inputs of sequence-type, blame and knowledge judgments, and all interactions showed that sequence-type was the only significant predictor ($F = 6.3084$; $p = .00448$); all other p-values were at least .28. Tukey HSD post hoc tests showed that the

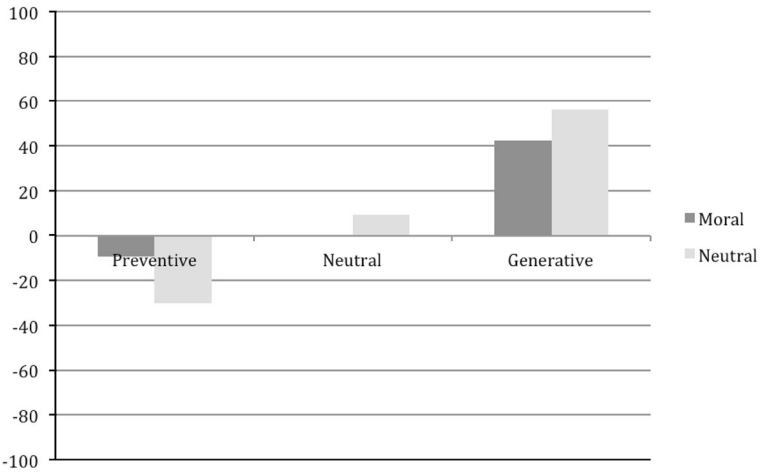


Fig. 2 Mean strength ratings in each condition

Generative sequence-type led to significantly different ratings than the Neutral and Preventive sequence-types in the restricted domain of only the Moral condition, though the latter two were not significantly different (G vs. P: $p = .0033$; G vs. N: $p = .0172$; N vs. P: $p = .818$).

An ANOVA of Blame judgments using sequence-type, knowledge judgments, strength ratings, and all interactions found only a main effect of sequence-type ($F = 3.6638$; $p = .036$; all other p-values $> .09$). Tukey HSD post hoc tests showed that this was due to Blame judgments for the Preventive sequences being significantly higher than those judgments for the Generative sequences ($p = .0096$). An ANOVA of Knowledge judgments using sequence-type, blame judgments, strength ratings, and all interactions found only a main effect of sequence-type ($F = 6.6906$; $p = .0034$; all other p-values $> .08$). Tukey HSD post hoc tests showed that Knowledge judgments for the Preventive sequences were significantly higher than for the Generative ($p = .00067$) and Neutral ($p = .0066$) sequences, though the latter two were not significantly different ($p = .72$).

Appendix 2

Participants

A total of 191 participants were recruited through Amazon's Mechanical Turk. The task took approximately 5 min to complete. Each participant was paid $.15 for participation.

Materials and methods

All participants were randomly assigned to one of the six Scenario-type X Outcome conditions. The global introduction for the Control scenario was:

> Johnson is a doctor traveling to the South Pacific Islands to research the rare skin disease *Anthrapora* that has been reported on various islands. In particular, Johnson is studying on the island of Tongatapu the impact (if any) of a native plant, *Solanaceae delisa*, on the skin disease *Anthrapora*. The plant may lead to the skin disease, it may cure the disease, or it may have no real effect at all. It is your job to figure this out.
>
> Johnson interviewed various villagers; some have the local disease, and some do not. She can diagnose villagers as suffering from the skin disease by finding the characteristic rashes. Unfortunately, because of language barriers, the only other information she can get from the villagers is whether or not they have come in contact with the native plant, *Solanaceae delisa*.

After reading the global introduction, participants were presented with one of the following outcomes:

(1) After observing 100 individuals, Johnson notes that most of the people who came into contact *Solanaceae delisa* had the skin disease *Anthrapora*, while a

few of the people who did *not* come into contact with *Solanaceae delisa* had the skin disease *Anthrapora*.

(2) After observing 100 individuals, Johnson notes that a few of the people who came into contact *Solanaceae delisa* had the skin disease *Anthrapora*, while most of the people who did *not* come into contact with *Solanaceae delisa* had the skin disease *Anthrapora*.

(3) After observing 100 individuals, Johnson notes that a few of the people who came into contact *Solanaceae delisa* had the skin disease *Anthrapora*, while most of the people who did *not* come into contact with *Solanaceae delisa* had the skin disease *Anthrapora*.

Participants were then given the following instructions:

> You will now be asked to evaluate the causal connection between the plant and the skin disease on a scale from −10 to +10. Respond with −10 if you think that exposure to the plant (*Solanaceae delisa*) *always prevents* the skin disease (*Anthrapora*). Respond with +10 if you think that exposure to the plant *always produces* the rashes. And respond 0 if you think the plant is *irrelevant* for whether the person suffers from the disease or is healthy. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

In each case, participants were asked "To what extent does coming into contact with the plant, *Solanaceae delisa*, cause the skin disease, *Anthrapora*?" Ratings were made on a scale ranging from −10 to 10.

The global introduction for the *Moralized* scenario was a follows:

> Smith is an elderly man who has devoted his life to cancer research. He has been involved in the development of various treatments, which have helped to save the lives of thousands of people. Recently, Smith traveled deep into the Amazon in order to recover a nearly extinct species of orchids called *Betrafindalis copernicia*.
>
> *Betrafindalis copernicia* contains a highly concentrated form of the chemical *dispofignila*. Smith has been experimenting with a synthetic form of *dispofignila* and has found that it slows down the growth of cancer. While synthetic doses slow the growth down somewhat, only the strongest form of *dispofignila*, which cannot be synthetically produced and is only found in the orchid *Betrafindalis copernicia*, slows down the growth substantially, almost to the point of stopping growth altogether. As a matter of fact, Smith was actually able to experiment with one *Betrafindalis copernicia*. He found that the *dispofignila* found in the *Betrafindalis copernicia* was in fact more potent than the synthetic form of *dispofignila* and that it lead to a significantly greater decrease in cancer growth than the synthetic form of *dispofignila*.
>
> *Betrafindalis copernicias* are very rare. They are only found in a remote part of the Amazon, and, because of global warming, only a few dozen plants survive even there. Smith is sure that if he can preserve these plants, then he

can develop a cure for cancer. Thus, Smith traveled to the Amazon, returned home safely with all of the orchids, and placed them in his greenhouse.

Smith's neighbor, Jones, hates Smith. Jones has always despised Smith for no good reason. Jones knows that Smith has recently returned with the only remaining orchids in the world and he wants to kill all of the plants and destroy Smith's hopes for finding a cure for cancer.

Jones wants to kill the orchids, but he doesn't want to get caught. He knows that he could simply uproot the plants and kill them, but in order to do that he would have to get into the greenhouse, which is secured by an alarm. However, there is one way that Jones can kill the plants without leaving any evidence. There are several hoses that run a steady flow of water into the greenhouse. Jones knows that if he can inject poison into the hoses, it will kill all of the plants, and no evidence will be left behind.

One night, Jones breaks into an old farmer's shed and finds several bottles with the label "poison"; Jones grabs one of the bottles and discretely leaves the shed. Unbeknownst to Jones, however, the farmer reuses his bottles: some of the ones labeled "poison" have fertilizers, others simply have water, and of course, some actually do have poison. (The farmer has a system for knowing what each bottle contains, though Jones obviously does not know this system.)

A few days later, Jones fills a syringe with the liquid from the stolen bottle. He then goes to Smith's greenhouse, pushes the needle through one of the hoses, and injects the liquid into that particular hose.

After reading the global introduction, participants were presented with one of the following outcomes:

(1) Smith enters the greenhouse the next day and does a check-up on the plants. After observing 100 plants, Smith notes that *most* of the plants that were watered by hose A (the hose that, unbeknownst to Smith, Jones injected the liquid into) were dead, while a *few* of the plants that were *not* watered by hose A were dead.
(2) Smith enters the greenhouse the next day and does a check-up on the plants. After observing 100 plants, Smith notes that a few of the plants that were watered by hose A (the hose that, unbeknownst to Smith, Jones injected the liquid into) were dead, while most of the plants that were *not* watered by hose A were dead.
(3) Smith enters the greenhouse the next day and does a check-up on the plants. After observing 100 plants, Smith notes that about half of the plants that were watered by hose A (the hose that, unbeknownst to Smith, Jones injected the liquid into) were dead, while about half of the plants that were *not* watered by hose A were dead.

Participants were then given the following instructions:

You will now be asked to evaluate the causal connection between the liquid and the plant dying on a scale from −10 to +10. Respond with −10 if you

think that exposure to the liquid *always kills* the plant. Respond with +10 if you think that exposure to the liquid *always prevents* the plant from dying. And respond 0 if you think the liquid is *irrelevant* for whether the plant dies or not. Please give your best estimate of the causal strength, even if you are uncertain about what is actually happening.

In each case, participants were asked "To what extent does coming into contact with the liquid cause the plant to die?" Ratings were made on a scale ranging from −10 to 10. Additionally people were asked "How blameworthy do you think Jones is for attempting to kill Smith's plants?" (9-pt scale anchored at 1 = "not at all blameworthy", 9 = "extremely blameworthy") and "Do you think that Jones knew what was in the bottle of liquid that he stole?" (9-pt scale anchored at 1 = "he did not know at all", 9 = "he definitely knew").

Results and analysis

We began by conducting an ANOVA with Condition and Outcome as predictors of causal ratings. We found a main effect of Condition ($F = 26.615$, $p = .000$) and Outcome ($F = 43.252$, $p = .000$) and no interaction effect ($F = 1.853$, $p = .160$). Importantly, we examined the effects of Condition on each Outcome. For the generative outcome i.e., "most", we found a significant effect of Condition ($F = 5.101$, $p = .027$) with people in the Moralized scenario ($M = 6.15$, $SD = 2.62$) making significantly higher causal ratings than those in the Neutral scenario ($M = 4.31$, $SD = 3.79$). For the preventative outcome i.e., "few", we, again, found a significant effect of Condition ($F = 5.102$, $p = .027$) with people in the Moralized scenario ($M = .500$, $SD = 5.21$) making significantly higher causal ratings than those in the Neutral scenario ($M = -1.90$, $SD = 3.11$). Finally, for the irrelevant outcome i.e., "half", we found a significant effect of Condition ($F = 19.544$, $p = .000$) with people in the Moralized scenario ($M = 2.60$, $SD = 3.62$) making much higher causal ratings than those in the Neutral scenario ($M = -1.77$, $SD = 4.08$).

References

Ahn, W.-K., & Bailenson, J. (1996). Causal attribution as a search for underlying mechanisms: An explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology, 31*, 82–123.

Alexander, J., Mallon, R., & Weinberg, J. M. (2010). Accentuate the negative. *Review of Philosophy and Psychology, 1*, 297–314.

Alicke, M. (1992). Culpable causation. *Journal of Personality and Social Psychology, 36*, 368–378.

Alicke, M. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*, 556–574.

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

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30*, 241–297.

Blakemore, S.-J., Fonlupt, P., Pachot-Clouard, M., Darmon, C., Boyer, P., Meltzoff, A. N., et al. (2001). How the brain perceives causality: An event-related fMRI study. *NeuroReport, 12*, 3741–3746.

Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking & Reasoning, 8*, 269–295.

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.

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439–477.

Cummins, D. D. (1996). Evidence of deontic reasoning in 3-and 4-year-old children. *Memory & Cognition, 24*(6), 823–829.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17*, 1082–1089.

Danks, D. (2009). The psychology of causal perception and reasoning. In H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *Oxford handbook of causation* (pp. 447–470). Oxford: Oxford University Press.

Danks, D., & Schwartz, S. (2005). Causal learning from biased sequences. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual meeting of the cognitive science society* (pp. 542–547). Mahwah, NJ: Lawrence Erlbaum Associates.

Danks, D., & Schwartz, S. (2006). Effects of causal strength on learning from biased sequences. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual meeting of the cognitive science society* (pp. 1180–1185). Mahwah, NJ: Lawrence Erlbaum Associates.

Driver, J. (2008a). Attributions of causation and moral responsibility. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 2, pp. 423–440)., The cognitive science of morality: Intuition and diversity Cambridge, MA: MIT Press.

Driver, J. (2008b). Kinds of norms and legal causation: Reply to Knobe and Fraser and Deigh. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 2, pp. 459–461)., The cognitive science of morality: Intuition and diversity Cambridge, MA: MIT Press.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S., Hau, R., et al. (2010). A choice prediction competition, for choices from experience and from description. *Journal of Behavioral Decision Making, 32*, 15–47.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255–278.

Feinberg, J. (1970). *Doing and deserving: Essays in the theory of responsibility*. Princeton: Princeton University Press.

Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.

Fugelsang, J. A., Roser, M. E., Corballis, P. M., Gazzaniga, M. S., & Dunbar, K. N. (2005). Brain mechanisms underlying perceptual causality. *Cognitive Brain Research, 24*, 41–47.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3–32.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834.

Hart, H. L. A., & Honoré, T. (1985). *Causation in the law* (2nd ed.). New York: Oxford University Press.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making, 32*, 48–68.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making, 21*, 493–518.

Heider, F., & Simmel, M.-A. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243–249.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*, 517–523.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *Journal of Philosophy, 106*, 587–612.

Knobe, J. (2009). Folk judgments of causation. *Studies in the History and Philosophy of Science, 40*, 238–242.

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences, 33*, 315–329.

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 2, pp. 441–448)., The cognitive science of morality: Intuition and diversity Cambridge, MA: MIT Press.

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

Machery, E. (2011). Thought experiments and philosophical knowledge. *Metaphilosophy, 4*, 191–214.

Machery, E. (2012). Power and negative results. *Philosophy of Science, 79*, 808–820.

Machery, E. (ms). *Evidence and Cognition*.

Macpherson, F. (2012) Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research, 84*(1), 24–62.

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

McGrath, S. (2005). Causation by omission: A dilemma. *Philosophical Studies, 123*, 125–148.

Michotte, A. (1946/1963). *The perception of causality*. London: Methuen.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language, 24*, 586–604.

Pinillos, N. Á., Smith, N., Nair, G. S., Marchetto, P., & Mun, C. (2011). Philosophy's new challenge: Experiments on intentional action. *Mind & Language, 26*, 115–139.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875–881.

Roser, M. E., Fugelsang, J. A., Dunbar, K. N., Corballis, P. M., & Gazzaniga, M. S. (2005). Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology, 19*, 591–602.

Satpute, A. B., Fenker, D. B., Waldmann, M. R., Tabibnia, G., Holyoak, K. J., & Lieberman, M. D. (2005). An fMRI study of causal judgments. *European Journal of Neuroscience, 22*, 1233–1238.

Schlottmann, A. (1999). Seeing it happen and knowing how it works: How children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology, 35*, 303–317.

Schlottmann, A. (2000). Is perception of causality modular? *Trends in Cognitive Sciences, 4*, 441–442.

Schlottmann, A., & Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *Quarterly Journal of Experimental Psychology, 44A*, 321–342.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*, 299–309.

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared make your tummy ache? Naive theories, ambiguous evidence and preschoolers' causal inferences. *Developmental Psychology, 43*, 1124–1139.

Sloman, S. (2009). *Causal models: How people think about the world and its alternatives*. Oxford: Oxford University Press.

Stokes, D. (2012). Perceiving and desiring: A new look at the cognitive penetrability of experience. *Philosophical Studies, 158*(3), 477–492.

Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition, 119*, 454–458.

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,* 814–820.

Taylor, R. S., & Chi, M. T. H. (2006). Simulation versus text: Acquisition of implicit and explicit information. *Journal of Educational Computing Research, 35*, 289–313.

Van Lehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences, 2*, 1–59.