COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 44 (2021) e12931

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12931

Norms Affect Prospective Causal Judgments

Paul Henne,[a,b] Kevin O'Neill,[c] Paul Bello,[d] Sangeet Khemlani,[d] Felipe De Brigard[c,e,f,g]

[a]*Department of Philosophy, Lake Forest College*
[b]*Neuroscience Program, Lake Forest College*
[c]*Center for Cognitive Neuroscience, Duke University*
[d]*Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory*
[e]*Department of Philosophy, Duke University*
[f]*Department of Psychology and Neuroscience, Duke University*
[g]*Duke Institute for Brain Sciences, Duke University*

Received 2 January 2020; received in revised form 8 October 2020; accepted 26 November 2020

Abstract

People more frequently select norm-violating factors, relative to norm-conforming ones, as the cause of some outcome. Until recently, this abnormal-selection effect has been studied using retrospective vignette-based paradigms. We use a novel set of video stimuli to investigate this effect for prospective causal judgments—that is, judgments about the cause of some future outcome. Four experiments show that people more frequently select norm-violating factors, relative to norm-conforming ones, as the cause of some future outcome. We show that the abnormal-selection effects are not primarily explained by the perception of agency (Experiment 4). We discuss these results in relation to recent efforts to model causal judgment.

Keywords: Causation; Causal selection; Causal reasoning; Causal judgment; Norms

1. Introduction

Suppose two cars crashed at an intersection; one went through a green light and the other through a red light. People might reason that the car's failure to stop at a red light caused the crash. A complex set of conditions, however, brought about the crash—one car went through the green light, for instance, and no obstacles came between the cars. Yet it is tempting to select only the car that went through the red light as *the* cause of the crash. Researchers refer to the difficulty in understanding why people select a

Correspondence should be sent to Paul Henne, Durand Art Institute, Lake Forest College, 555 North Sheridan Road, Lake Forest, IL 60045. E-mail: phenne@lakeforest.edu

particular factor as *the* cause over all the other necessary factors as *the causal selection problem* (Hart & Honore, 1985; Hesslow, 1988).

Until recently, theorists dismissed the selection problem as an issue of pragmatics, irrelevant to understanding the meaning of causal statements. David Lewis famously wrote that he had “nothing to say about these principles of invidious discrimination” that distinguish necessary causal factors and “the” cause (Lewis, 1974). And, following Lewis, many philosophers have argued that causal selection has nothing to do with the nature of causation itself or the semantics of causal language—rather, it is some capricious feature of conversational pragmatics (for discussion, see Driver, 2008; Schaffer, 2013). However, a number of cognitive scientists and philosophers have recently turned their attention to the causal selection problem in order to explain why people select some causes over others (Bernstein, 2014, 2015; Henne, Pinillos, & De Brigard, 2017; Hitchcock & Knobe, 2009; McGrath, 2005).

Some of the researchers who attempt to explain the causal selection problem argue that norms affect causal selection such that people select norm-violating factors as causes (Hart & Honore, 1985). Norms can encompass statistical regularities, prescriptions for behavior, and intended functions (Bear & Knobe, 2017; Kominsky & Phillips, 2019). On such views, the car that went through the red light violated a prescriptive norm—that is, cars should stop at red lights—so people select it relative to the other norm-conforming factors—for example, the car that went through the green light. And extensive empirical work shows that, in fact, people do tend to select norm-violating rather than norm-conforming factors as the cause (Gerstenberg & Icard, 2019; Hilton & Slugoski, 1986; Hitchcock & Knobe, 2009; Icard, Kominsky, & Knobe, 2017; Kirfel & Lagnado, 2018; Knobe & Fraser, 2008; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Morris, Phillips, Gerstenberg, & Cushman, 2019), both in the case of actions (Phillips, Luguri, & Knobe, 2015) and in the case of omissions (Henne et al., 2017). Recent computational and formal frameworks of causal reasoning represent norms directly to account for this effect (Bello, 2016; Halpern & Hitchcock, 2014; Icard et al., 2017).

There are several explanations for why people select norm-violating factors as causes. Some theorists argue that these effects result from people’s tendency to blame agents (Alicke, 1992; Alicke, Mandel, Hilton, Gerstenberg, & Lagnado, 2015; Alicke, Rose, & Bloom, 2011; Rogers et al., 2019; Shaver, 1985). On these views, people’s desire to blame a person biases their causal judgments (Alicke, 1992; Rogers et al., 2019). Someone who violates a norm is more blameworthy than someone who does not, and this desire to blame such an individual makes people more inclined to identify the actions of that individual as the cause of the outcome. These theories suggest that, for instance, people would judge that the car that went through the red light caused the crash because people align their causal judgment with their desire to blame the driver of the car that went through the red light. These views hold that prescriptive norms alone could not explain abnormal-selection effects, as social reactions to agents such as blame and praise drive causal attributions (Alicke et al., 2011, 2015). Hence, abnormal-selection effects are best explained by features of social cognition. We refer to these accounts as *social-cognition explanations*.

There is some evidence in favor of social-cognition explanations. In one study, researchers gave participants vignettes describing intentional and unintentional actions and their consequences, and they found that participants' causal judgments and blame judgments were highly correlated and that both kinds of judgments were affected by the manipulations in similar ways (Lagnado & Channon, 2008). Critics may counter that evaluations of blame are often retrospective and can depend on the consequences of the action; for instance, people receive lesser charges for an attempted murder than for an actual murder. As such, social-cognition explanations may have difficulty explaining how norms affect anticipated causal judgments for outcomes that have yet to occur.[1]

Other researchers argue that abnormal-selection effects reflect more general cognitive processes than social cognition (Gerstenberg & Icard, 2019; Henne, Bello, Khemlani, & De Brigard, 2019; Hitchcock & Knobe, 2009; Icard et al., 2017). These researchers argue that norms affect people's explicit (Kominsky & Phillips, 2019; Phillips et al., 2015) or implicit (Icard et al., 2017) consideration of relevant alternatives (see also Henne, Niemi, Pinillos, De Brigard, & Knobe, 2019; Morris et al., 2019), and because thinking about alternatives affects causal judgments (Byrne, 2016), norms, at least in many cases, affect causal judgments. On such views, people are more inclined to consider the counterfactuals where the one car did not go through the red light (i.e., what should have happened), so they tend to agree that the car going through the red light caused the crash. Some others in this group argue that the consideration of particular alternatives determines causal judgments (Johnson-Laird & Khemlani, 2017; Khemlani, Wasylyshyn, Briggs, & Bello, 2018) and that norms help people to construct those alternatives (Bello & Khemlani, 2015; Henne, Niemi, et al., 2019). While these views vary widely, we will generally refer to them as *modal explanations*, since they argue that norms affect the possibilities people consider.

There is some evidence in favor of modal explanations. In one study, researchers gave participants vignettes describing agents acting immorally or inanimate objects malfunctioning (Kominsky & Phillips, 2019: Experiment 4). Participants in this study responded similarly to both types of vignettes: They tended to judge the norm-violating factor as the cause, regardless of whether the factor was an agent acting immorally or an object malfunctioning (Kominsky & Phillips, 2019: Experiment 4). Since only agents can be blamed—agents carry out volitional, intentional actions, while inanimate objects do not—the authors took the result to suggest that more general cognitive processes were at play (Kominsky & Phillips, 2019). Verb usage can make inanimate objects seem like agents, so critics might argue that descriptions of inanimate objects are inherently agentive (see Rose, 2017). But whether this tendency is pervasive is an ongoing debate. Notably, modal explanations can easily accommodate both retrospective and prospective causal judgments; if norms affect people's consideration of retrospective possibilities, they can affect people's consideration of prospective ones, too.

In the present paper, we have two primary aims. First, we sought to test whether norms affect prospective causal judgments—that is, judgments about the cause of some future outcome when a set of potential causal factors have already occurred. Reasoners may select a normal-violating factor as the cause of a future outcome, just as they do for a known outcome. Only modal explanations, however, predict that they should. On social cognition views, people align their causal judgments with their desire to blame the source of their reactions (Alicke, 1992; Rogers et al., 2019). So, if the undesirable outcome has not yet occurred, then there is no reaction that motivates the alignment of a causal judgment with their desire to blame. If blame validation explains abnormal-selection effects, then these effects should not arise for prospective causal judgments until the outcome is known. Some social-cognition explanations might be extended to cover judgments about future effects and anticipated blame. Since some views hold that people align their causal judgment with their *desire* to blame or praise the source of their reactions, they may not need to rely on the judgments being retrospective (Alicke et al., 2011). If people exhibit consistent abnormal-selection effects for prospective causes, then social-cognition explanations need to be revised to accommodate them. Some of the modal explanations, however, clearly predict that norms should affect prospective causal judgments just as they do for retrospective causal judgments (e.g., Byrne, 2016). If people simulate possible alternatives to potential causal factors when making their causal judgments, then they should be doing the same when thinking about prospective causes. Suppose the two cars had gone through the red light and the green light but the crash had yet to occur. Some modal views suggest that people would be more inclined to consider alternatives where the car did not go through the red light, making them more inclined to state that the car going through the red light would cause the crash. But no studies reveal whether this tendency happens for prospective events (Byrne, 2016). Determining whether abnormal-selection effects occur for prospective causal judgments will help decide between the social-cognition and modal explanations.

Second, we sought to obtain people's causal judgments of nonsocial, visual stimuli—that is, videos of physical objects interacting. Most studies exploring abnormal-selection effects have used verbal vignettes to prompt causal judgments (e.g., Henne et al., 2017; but see Gerstenberg & Icard, 2019; Kirfel & Lagnado, 2018), which require participants to imagine complex interactions between different entities and to base their judgments on what they imagine. But in vignette-based studies, participants have little guidance on what to imagine when they mentally simulate the scenarios described in the vignettes. The descriptions of the scenarios may make people perceive non-agentive scenarios as agentive (Rose, 2017). By using imagery of interacting physical objects, we should reduce the opportunity for participants to view objects as agents. Social-cognition explanations would predict that people's causal judgments should be less prone to abnormal-selection effects when the stimuli do not concern agents, since only agents can be blamed. But if people's causal judgments show abnormal-selection effects in the absence of any opportunity to blame an agent, the result would support modal explanations of causation.

An additional benefit of using visual stimuli is that it allows for participants to respond to many different variations of the investigated interactions. Most studies that investigate abnormal-selection effects use only a single verbal vignette (e.g., Henne et al., 2017; but see Icard et al., 2017). Using such a small number of stimuli in experiments limits the generalizability of the results and allows for the possibility that earlier findings resulted from peculiar experimental stimuli. Our studies accordingly use multiple videos with different configurations to test the predictions of the competing theories.

We report four experiments below that reveal abnormal-selection effects for prospective causal judgments. In Experiment 1, we found that norms affect causal selection; for multiple videos, participants were more inclined to select norm-violating factors, relative to norm-conforming ones, as the cause of a future outcome. Experiment 2 replicated the pattern of results from Experiment 1 and extended it to novel stimuli. But, despite a robust abnormal-selection effect in aggregate, the effect varied across the different videos. Experiment 3 showed that the variance was likely due to a recency effect, and it too revealed a robust abnormal-selection effect. In Experiment 4, we manipulated the agency of the stimuli directly and asked participants about the intentionality and responsibility of the stimuli. Abnormal-selection effects did not differ as a function of perceived agency, supporting the modal explanation of these effects.

2. Experiment 1

Experiment 1 tested two predictions of the modal explanation for abnormal-selection effects. First, it examined whether abnormal-selection effects arise for prospective causal judgments. Second, it examined whether prescriptive norms affect participant's selection of causes in visual stimuli that are nonsocial. Specifically, it investigated whether participants are more likely to select norm-violating factors as a future cause of an outcome than they are to select those factors where no norm is specified.

To these ends, Experiment 1 showed participants three videos of different object configurations in which balls interact and move toward a goal. In each of the videos, two balls (A and B) collide and send one of the balls in a new direction. And in each of the videos, there is a paddle that can redirect the movement of the ball. We added a forked tube to each video to increase the variability of the interactions in each video. We manipulated whether the collision of the balls or the paddle's movement violated a norm. For instance, on some trials, participants were told that the collision of the balls was not supposed to occur. Critically, each video ended before ball A entered a goal, and participants were asked to select from a list of options which factor *will* cause the ball to enter the goal. The modal explanation predicts that norm-violating factors should be selected as the cause more often than norm-conforming factors.

2.1. Methods

2.1.1. Participants

We calculated the effect size for a GLM on a single item and then multiplied by 3 in order to calculate the sample size required for all three of our items. We used the pwr package (Champely et al., 2018) in R to conduct a power analysis for a single item in our study. Since our goal was to obtain .85 power to detect a medium-large effect ($f = 0.25$) at

.001 α error probability, 90 participants were required for each item in our study. So our target sample size was 270 participants. We anticipated a 5% dropout rate, so we recruited 285 participants through Amazon Mechanical Turk (AMT) for \$1.50 compensation to participate. We recruited participants with a 99% approval rating who were located in the United States and had not participated in a pilot experiment. We excluded 10 participants who failed to pay attention throughout the experiment, so we analyzed the data from 275 remaining participants ($M_{age} = 35$, $Range_{age} = [19–74]$, 35% female).

2.1.2. Preregistration and open science

The effects of Experiment 1 were preregistered through the Open Science Framework. All materials, preregistrations, analysis code, and data for Experiment 1 and all subsequent studies are available at https://osf.io/4pvyd/.

2.1.3. Design, materials, and procedure

The experiment first familiarized participants with the videos they were to watch in the experimental conditions and with the four elements in each video: the balls, the paddle, the goal, and a tube. The experiment further familiarized participants with the components by having them watch three practice videos. Specifically, participants watched a video where the two balls collided, and then the paddle blocked ball A from entering the tube; a video where the two balls collided, the paddle did not block ball A from entering the tube, and then ball A entered the goal; and then a video where the two balls collided, the paddle did not block ball A from entering the tube, and then ball A did not enter the goal. Each video was accompanied by a description of what occurred. The experiment then told participants that the instruction phase was over and that the test phase would start. The experiment assigned participants to all six study conditions in randomized order in a 3 (Item: V1, V2, V3) × 2 (Norm: Balls or Paddle) within-participants design (Fig. 1). Participants evaluated a total of six separate videos, which each ranged between 15 and 18 s and were presented in randomized order. For each trial, the experiment told participants that the collision was norm-violating (e.g., “the balls are **not** supposed to collide”) or else that the paddle’s movement was norm-violating (e.g., for V2: “the paddle is **not** supposed to block the tube”; and for V3: “the paddle is supposed to block the tube”). To verify that participants understood what the norm was, the experiment asked on the following screen about the norm (e.g., “In the video you’re about to watch, is the paddle supposed to block the tube?”). If participants answered correctly, they received confirmatory feedback and then were reminded of the norm. If they answered incorrectly, they received corrective feedback and were reminded of the norm. The experiment then reminded participants of the norm manipulation (e.g., “**Remember**: In the video you’re about to watch the paddle is supposed to block the tube”) and presented a video. For each video, participants responded to the following prompt:

Please select the phrase that best completes the following statement about the video you just watched: ________ will cause the ball to reach the goal.

In all conditions, participants selected a response from a drop-down menu that listed the following factors in randomized order: the balls colliding, the paddle moving (or not

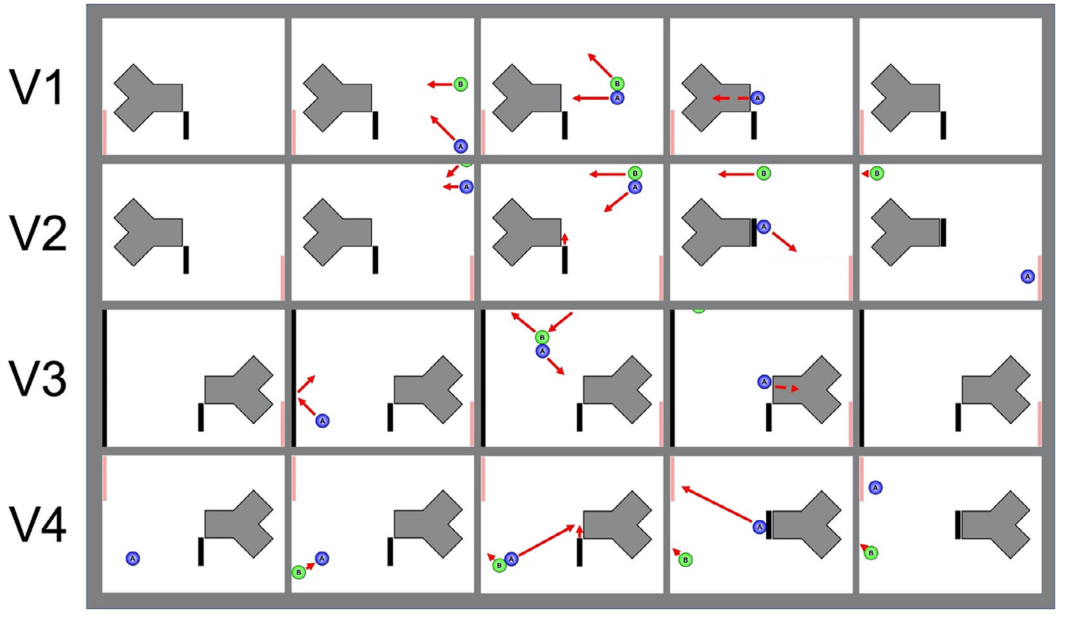

Fig. 1. Videos used in Experiment 1 (V1–V3) and Experiment 2 (V1–V4). In Experiment 2, V1 and V2 were oriented to match the direction of V3 and V4 so that the tube always had a left-to-right orientation. In each video, a blue ball (A) collided with a green ball (B) and moved in the direction of a gray forked tube. In some videos, the tube was blocked by the black paddle. The pink bar depicts the goal. Red arrows depict the direction in which the balls and the gate moved.

moving), the tube, or some other event. The wording for each response varied slightly for each video. For instance, the options for V1 were as follows: “The paddle’s failure to move,” “The collision of the balls,” “The forked shape of the tube,” and “Some other event.” After participants completed this task for all six videos, the experiment asked basic demographic information. The experiment then asked participants to respond to an explicit attention check that was used in all experiments (https://osf.io/kfj7u/).

2.2. Results

To test whether participants were more likely to select norm-violating factors as the cause of a future outcome, we first examined the difference in the selection of the potential causal factors as a function of the norm condition across all three videos.[2] In line with work on retrospective causal judgments, participants were more inclined to select norm-violating factors as the cause across all videos ($b = 1.66$, $SE = 0.11$, $z = 14.6$, $p < .001$, CI [1.43, 1.88]) and for each video (V1: $b = 2.65$, $SE = 0.23$, $z = 11.20$, $p < .001$, CI [2.18, 3.11]; V2: $b = 0.97$, $SE = 0.21$, $z = 4.51$, $p < .001$, CI [.55, 1.40]; V3: $b = 2.25$, $SE = 0.22$, $z = 9.80$, $p < .001$, CI [1.79, 2.69]) (Fig. 2A). We report the percentage of participants who selected each factor in each video in Table S1 and in Fig. S1A (https://osf.io/rzmv3/).

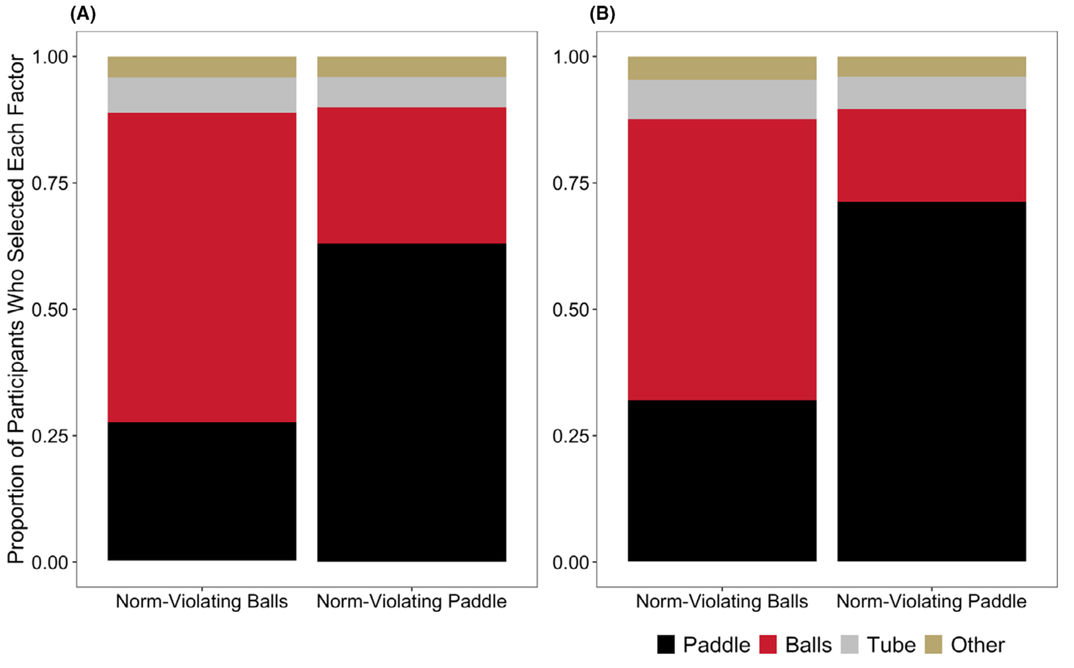

Fig. 2. Proportion of participants who selected each factor as the cause as a function of condition collapsed across all videos in Experiment 1 (A) and Experiment 2 (B).

2.3. Discussion

Participants more frequently selected norm-violating factors than factors that did not violate a norm as the cause of some future outcome. This pattern arose across multiple videos with visual stimuli that are nonsocial and that do not include agents. It suggests that perceived norms help people select between causes and causal conditions, and it is consistent with the modal explanation; in the absence of social information, participants showed an abnormal-selection effect.

There appeared to be a difference in participants' responses between one item (V2) and the others (V1 and V3) (Fig. S1A: https://osf.io/rzmv3/). For V2, participants selected the paddle's movement as the cause more often than the balls colliding, the tube, or something else altogether—even when the collision was norm-violating. While there was an abnormal-selection effect in each video, this effect was larger in V1 ($OR = 14.13$) and V3 ($OR = 9.47$) than in V2 ($OR = 2.66$).

One potential explanation for this unexpected result is that participants only saw the paddle move two times in the experiment (once during the instruction phase and once during V2). So participants could have perceived the movement of the paddle as statistically abnormal and the failure of the paddle to move as statistically normal (see Gerstenberg & Icard, 2019). We controlled for this confound in Experiment 2.

3. Experiment 2

Experiment 2 controlled for the statistical normality of the paddle moving or not moving by adding a fourth video (V4) where the paddle also moved. By adding this video, participants saw the paddle move or not move the same number of times in the experiment. We also controlled for the direction of the tube in the videos: All were oriented for left-to-right movement. Hence, in this experiment we used four videos in which two balls (A and B) collided and where a paddle moved to block a tube or did not move to block a tube. As in Experiment 1, all videos ended before ball A entered a goal, and we asked participants to select the factor that would cause ball A to enter the goal. For each of the four videos, the experiment varied the abnormality of the balls colliding or the paddle moving in the same manner as Experiment 1.

3.1. Methods

3.1.1. Participants

We kept the same sample size as that in Experiment 1 such that 90 participants were required for each item in our study. Hence, our target sample size was 360 participants. We anticipated a 5% dropout rate, so we recruited 380 participants through AMT for $2.50 compensation to participate. We recruited participants with a 99% approval rating who were located in the United States and had not participated in Experiment 1 or a pilot experiment. We excluded 21 participants who failed to pay attention throughout the experiment, so we analyzed the data from 359 remaining participants ($M_{age} = 32$, $Range_{age} = [18\text{-}64]$, 46% female).

3.1.2. Design, materials, and procedure

The experiment first familiarized each participant with the videos and the four components of the videos just as in Experiment 1. The experiment then told participants that the instruction phase was over and that the test phase would start. The experiment assigned participants to all eight conditions in randomized order in a 4 (Item: V1, V2, V3, V4) × 2 (Norm: Balls or Paddle) within-participants design (Fig. 1). The experimental design was exactly the same as that in Experiment 1. The only difference in this experiment was the additional, modified materials. After participants completed this task for all eight videos, the experiment asked them for basic demographic information. The experiment then asked participants to respond to the same explicit attention check used in Experiment 1.

3.2. Results

Participants most often selected norm-violating factors as the cause across all videos ($b = 1.91$, $SE = 0.09$, $z = 21.1$, $p < .001$, CI [1.73, 2.09]) and for each video (V1: $b = 2.61$, $SE = 0.20$, $z = 13.00$, $p < .001$, CI [2.21, 3.00]; V2: $b = 1.80$, $SE = 0.20$, $z = 8.86$, $p < .001$, CI [1.40, 2.20]; V3: $b = 2.61$, $SE = 0.20$, $z = 12.50$, $p < .001$, CI

[2.19, 3.01]; V4: $b = 1.81$, $SE = 0.20$, $z = 8.65$, $p < .001$, CI [1.39,2.21]) (Fig. 2B). We report the percentage of participants who selected each factor in each video in Table S2 and in Fig. S1B (https://osf.io/rzmv3/).

3.3. Discussion

Experiment 2 replicated the results of Experiment 1: Participants selected norm-violating factors as the future cause more often than other factors. That is, abnormal-selection effects arose across multiple, nonsocial visual stimuli for prospective causal judgments. This evidence supports the view that perceived norms helped participants select between causes and causal conditions, and it is consistent with the modal explanation for abnormal-selection effects.

Despite our additional controls for statistical normality and object orientation, there remained an apparent difference in participants' response between two items (V2 and V4) and the others (V1 and V3) (Fig. S1B: https://osf.io/rzmv3/). The proportion of participants selecting the paddle's movement as the cause in V2 and V4 was much higher than the proportion of participants selecting the paddle's failure to move as the cause in V1 and V3. While there was an abnormal-selection effect in each video, this effect was larger in V1 ($OR = 13.60$) and V3 ($OR = 13.58$) than in V2 ($OR = 6.05$) and V4 ($OR = 6.08$). Given that we controlled for the statistical normality of the paddle moving—how frequently participants saw it move or not move—we have no evidence that this control made a difference to people's responses. Thus, while the abnormal-selection effects predicted by the modal explanations is apparent in all items (V1, V2, V3, and V4), it was still surprising that the proportion of responses was different between V1 and V3 relative to V2 and V4.

One possible explanation for this difference in selection across items is that a recency effect could have also affected participants' judgments in our materials. After all, in V2 and V4 the paddle moves *after* the balls collide, whereas in V1 and V3 the paddle does not move at all. Several studies show that people are more inclined to select the most recent factor as the cause of the outcome (e.g., Lagnado & Channon, 2008; Reuter, Kirfel, Riel, & Barlassina, 2014; Spellman, 1997). Hence, there could have been interference in our results such that a recency effect in V2 and V4 made people more inclined to select the most recent factor—the moving of the paddle—rather than the abnormal factor. Experiment 3 investigated this possibility.

4. Experiment 3

In Experiment 2, we found that participants were overall more inclined to select the paddle's movement as the future cause of the outcome in some of our materials (V2 and V4). In these materials, the paddle's movement was the last event to occur before the end of the video. Because people are more inclined to select recent events as causes (e.g., Byrne, 2016; Reuter et al., 2014), a recency effect may explain the difference between the videos in Experiments 1 and 2. In Experiment 3, we tested whether the pattern occurred because of the temporal order of the factors.

To test this hypothesis, we modified V2 and V4 to manipulate the temporal order of potential causal factors. Specifically, for each video, we varied whether the balls collided first or the paddle moved first. We again varied whether the balls colliding or the paddle moving or not moving was norm-violating. As such, the modal explanation's prediction about norms applies in Experiment 3 as well: If norms affect the selection of a potential causal factor, then participants will select the norm-violating factor more frequently than any others. Specifically, we predicted that, in situations where the paddle moves first and the balls' collision is norm-violating, participants should select the balls more frequently as the cause. Moreover, if temporal order affects the selection of a potential causal factor, then participants should select the more recently occurring factor more frequently than any others.

4.1. Methods

4.1.1. Participants

We kept the same sample size as that in Experiment 2. As such, we recruited participants from AMT with a 99% approval rating who were located in the United States and had not participated in any other experiments in this article. A total of 382 such participants completed the survey for \$2.75 compensation to participate. We excluded 17 participants who failed to pay attention throughout the experiment, so we analyzed the data from 365 remaining participants ($M_{\text{age}} = 31$, $\text{Range}_{\text{age}} = [18–66]$, 50% female).

4.1.2. Materials, design, and procedure

The experiment first familiarized each participant with the videos and the four components of the videos just as in Experiment 1. The experiment then told participants that the instruction phase was over and that the test phase would start. The experiment assigned participants to all eight conditions in randomized order in a 2 (Item: V2 or V4) × 2 (Norm: Balls or Paddle) × 2 (Order: Balls First or Paddle First) within-participants design (Fig. 3). The experiment manipulated norms exactly the same as it did in Experiment 2. The only difference was that we manipulated V2 and V4 from Experiment 2 so that either the paddle moved first or the balls collided first. Hence, participants watched the same videos from Experiment 2 (V2 and V4) and the new manipulated versions. All videos lasted 15 s. After participants completed this task for all eight videos, the experiment asked them for basic demographic information. Participants then responded to the same explicit attention check used in Experiment 2.

4.2. Results

There was an interaction of norms and temporal order across all videos ($b = 0.04$, $SE = 0.16$, $z = 0.27$, $p < .001$, CI [−.28, .37]) and for each video (V2: $b = 0.21$, $SE = 0.23$, $z = 0.92$, $p < .001$, CI [−.24, .67]; V4: $b = -0.12$, $SE = .23$, $z = -0.52$, $p < .001$, CI [−.59, .34]). Participants selected the norm-violating factor as the cause more

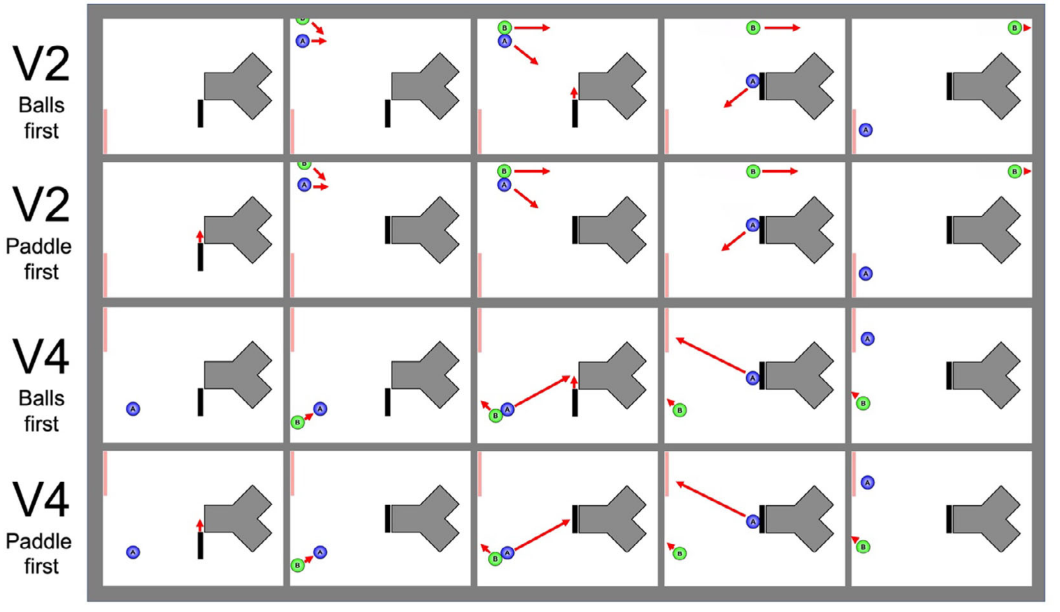

Fig. 3. Four videos used in Experiment 3. In each video, a blue ball (A) collided with a green ball (B) and moved in the direction of a gray forked tube. In some videos, the tube was blocked by the black paddle. The red arrows depict the motion of the balls, and the pink bar represents the goal.

frequently than they selected the factor that did not violate the norm across all videos ($b = 0.92$, $SE = 0.12$, $z = 7.52$, $p < .001$, CI [.68, 1.15]) and for each video (V2: $b = 0.81$, $SE = 0.17$, $z = 4.75$, $p < .001$, CI [.47, 1.14]; V4: $b = 1.03$, $SE = 0.17$, $z = 5.88$, $p < .001$, CI [.68, 1.37]). People also selected more recent factors as the cause across all videos ($b = -.57$, $SE = 0.11$, $z = -5.18$, $p < .001$, CI [−.78, −.35]) and for each video (V2: $b = -0.68$, $SE = 0.15$, $z = -4.34$, $p < .001$, CI [−.98, −.37]; V4: $b = -0.46$, $SE = 0.15$, $z = -2.99$, $p < .001$, CI [−.76, −.15]) (Fig. 4).

4.3. Discussion

Consistent with Experiments 1 and 2, participants in Experiment 3 selected norm-violating factors as the future cause of an outcome more often than they selected other factors. This response pattern arose across multiple items with visual stimuli that are nonsocial and that do not include agents. Experiment 3 further revealed a recency effect: Participants were more inclined to select the most recent factor as the cause of the outcome. These finding suggests that recency may explain the differences in causal-selection behavior between items in Experiment 1 and 2. Future work should further investigate whether uncertainty (Spellman, 1997) or action–omission difference (Henne, Niemi, et al., 2019) explains these apparent differences in prospective causal judgments.

A criticism of Experiments 1–3 is that, despite their use of visual and nonsocial stimuli, the experiments may not have eliminated agency cues entirely. After all, people can

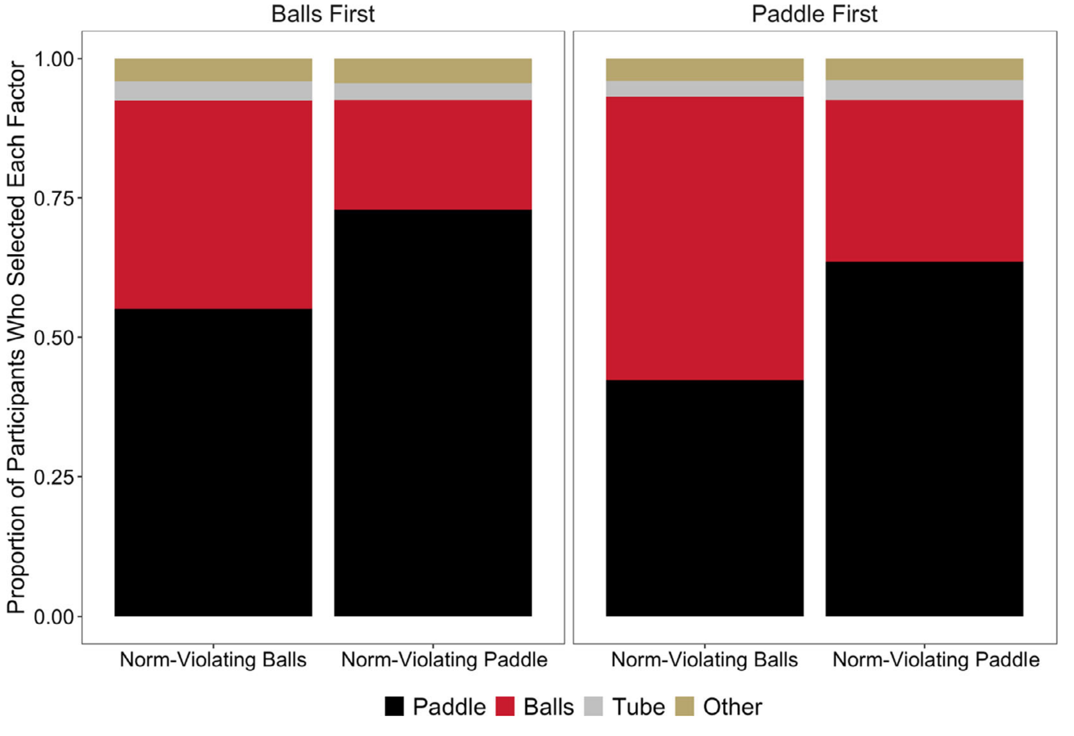

Fig. 4. Proportion of participants selecting each factor as the cause as a function of condition collapsed across all videos in Experiment 3.

perceive basic geometric figures as agents engaging in social interactions when those agents make movements that mimic emotional states (e.g., Heider & Simmel, 1944). In Experiments 1–3, we did not measure participants' perception of social or agentive aspects of the stimuli, so we cannot determine the extent to which people interpreted the materials as social. Moreover, Experiments 1–3 used arbitrary prescriptive norms which could prompt participants to think about the components in the videos as agentive. When the experiments told participants that the paddle should not move, for instance, participants may have attributed intentionality or other agentive features to the non-agentive components in the videos. We address these concerns in Experiment 4.

5. Experiment 4

We ran Experiment 4 to determine if perceived agency affects abnormal-selection behavior, as the social views would suggest. As in previous experiments, Experiment 4 varied whether the collision of the balls or the paddle's lack of movement violated a norm. Unlike previous experiments, however, this experiment established the norm statistically rather than prescriptively. Before the test video, half the participants watched five videos where the balls did not collide—hence, the subsequent collision in the test video violated a statistical norm. The other half of the participants watched five videos where the paddle moved to block the ball from entering the tube—hence, a subsequent video in which the paddle failed to movement violated a statistical norm.

The experiment also varied the agentive aspects of the materials both visually and descriptively. Half of the materials used balls that were visually similar to those used in the previous experiments, and the experiment stated that the balls are physical objects and that their movements are determined by physics. The other half of the materials used smiley-face icons in the place of the balls, and the experiment stated that the balls are being controlled by humans. To assess the effectiveness of such a manipulation, the study measured participants' perception of intentionality and responsibility for the balls and the paddle. If the modal explanation is correct, then the agency manipulation should affect people's judgments of intentionality and responsibility but not their causal-selection behavior; rather, causal-selection behavior should only vary as a function of norms.

5.1. Methods

5.1.1. Participants

We used the pwr package (Champely et al., 2018) in R to conduct a power analysis to obtain 0.9 power to detect a small-medium effect ($f = 0.1$) at 0.001 α error probability. We required a total of 212 participants. We anticipated a 1% dropout rate, so we recruited 215 participants through *Prolific*. All participants were U.S. nationals, were born in and resided in the United States, spoke English as their first language, and had a 99% approval rating on *Prolific*. A total of 214 such participants completed the survey for $1.00 compensation to participate. We excluded three participants who failed to pay attention throughout the experiment, so we analyzed the data from 211 remaining participants ($M_{\text{age}} = 33$, $\text{Range}_{\text{age}} = [18–74]$, 52% female).

5.1.2. Materials, design, and procedure

The experiment first familiarized participants with the videos and the four components of the videos just as in the previous experiments. The only difference was that the object orientations in this experiment were slightly different (Fig. 5). The experiment randomly assigned participants to one of four conditions in a 2 (Norm: Balls or Paddle) × 2 (Agency: High or Low) between-participants design. Participants read instructions specific to the agency manipulation to which they were assigned. Participants in the low-agency condition read: "In the videos, balls will move across the screen and interact with these components. These balls are simply mechanical objects. Their movements are completely determined by the laws of physics." Participants in the high-agency condition read: "In the videos, balls will move across the screen and interact with these components. These balls, unlike the other objects, are digital avatars controlled by people. Their movements are consistent with the laws of physics." The experiment then asked participants in both conditions, "Can the balls move of their own accord?," and they selected either "Yes, they are

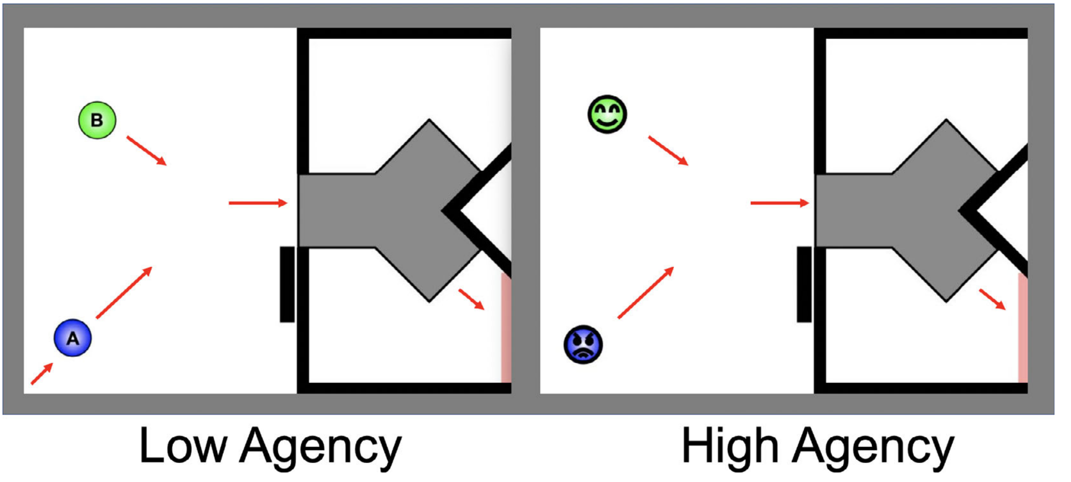

Fig. 5. Test videos used in Experiment 4. In the low-agency conditions, ball A enters from off screen and collides with ball B, and then ball B enters the gray forked tube before exiting the tube downward toward the goal. In the high-agency conditions, the blue ball self-propels toward the green ball and collides with it, and then the green ball enters the gray forked tube before exiting the tube downward toward the goal.

controlled by people" or "No, they are inanimate objects." The experiment corrected participants and reminded them of the instruction if they got this question incorrect. The experiment reminded participants of the instruction if they got this question correct. The experiment also emphasized this manipulation visually in all videos that followed. In the low-agency condition, ball A started its movement off screen and the balls were visually similar to those in the previous experiments. In the high-agency condition, the blue ball (the analog to ball A in the low-agency conditions) started moving on screen so that its movement appeared to be self-propelled, and the balls had faces on them to remind participants of the agency manipulation. Participants then entered the statistical norm-learning phase. On the next screen, participants read, "You will now be asked to watch five videos. After each video, you will be asked to determine if the balls collided and if the paddle moved in each individual video." Participants then watched five 15-s videos in which the balls moved from starting places and at different angles, and they reported whether the balls collided or the paddle moved on a binary (Yes/No) selection. In the norm-violating-balls condition, the balls did not collide and the paddle did not move in any of the five videos. As such, it was statistically normal for the balls not to collide. In the norm-violating-paddle conditions, the balls collided and the paddle moved to block the balls in all five videos. As such, it was statistically normal for the paddle to move and block the ball. After participants watched all five videos, the experiment asked, "Given the videos you've seen, how likely is it for balls to collide?" and "Given the videos you've seen, how likely is it for the paddle to move?" Participants responded to each manipulation-check question on a 0% likely to 100% likely scale. Participants were then told that they were finished with the instructions and that they would be asked to watch one more video.

In the test video, the balls collided, the paddle did not move, and ball B (or the blue ball in the high-agency condition) approached the goal (Fig. 5). The video stopped before it reached the goal. Below the video, participants then read the following prompt:

Please select the phrase that best completes the following statement about the video you just watched: ________ will cause [ball B/the green ball] to reach the goal.

Participants selected a response from a drop-down menu that listed the following factors: "The paddle's lack of movement," "The collision of the balls," "The forked shape of the tube," and "Some other event." On the following page, participants then answered four questions about the intentionality and responsibility of the ball and the paddle on a −50 to 50 scale [−50 = not at all, 50 = totally]: "To what extent do you believe that Ball A [the blue ball] **intentionally** collided with Ball B [the green ball]?"; "To what extent do you believe that if Ball B [the green ball] enters the goal, Ball A [the blue ball] is responsible?"; "To what extent do you believe that the paddle **intentionally** did not move to block Ball B [the green ball] from entering the tube?"; and "To what extent do you believe that if Ball B [the green ball] enters the goal, the paddle is responsible?" After participants completed this task, the experiment asked for basic demographic information. Participants then responded to the same explicit attention check used in all other experiments.

5.2. Results

Our manipulation checks suggested that the statistical norm manipulations were effective (Fig. S2: https://osf.io/rzmv3/). People in the norm-violating-balls conditions overwhelmingly thought that the collision of the balls was unlikely ($M = 18.42$, $SD = 22.78$, $n = 102$), relative to those in the norm-violating-paddle conditions ($M = 97.18$, $SD = 6.48$, $n = 109$). People in the norm-violating-paddle conditions overwhelmingly thought that the paddle's movement was likely ($M = 94.90$, $SD = 12.65$, $n = 109$), relative to those in the norm-violating-balls conditions ($M = 7.56$, $SD = 13.04$, $n = 102$).

5.2.1. Intentionality attributed to ball

Participants in the high-agency condition were more inclined to say that the blue ball intentionally collided with the green ball than participants in the low-agency condition were to say that ball A intentionally collided with ball B ($F(1, 207) = 52.95$, $p < .001$, $\eta_p^2 = .2$, CI [.13, .28]) (Fig. 6A). There was no main effect of norm ($F(1, 207) = 1.94$, $p = .16$, $\eta_p^2 = .01$, CI [.00, .04]), and there was no interaction ($F(1, 207) = 1.94$, $p = .69$, $\eta_p^2 = .00$, CI [.00, .02]).

5.2.2. Responsibility attributed to ball

We had no evidence that participants in the high-agency condition were more inclined to say that the blue ball would be responsible for the green ball entering the goal than participants in the low-agency condition were to say that ball A would be responsible for ball B entering the goal ($F(1, 207) = 1.93$, $p = .16$, $\eta_p^2 = .01$, CI [.00, .04]) (Fig. 6B). There was also no main effect of norm ($F(1, 207) = .91$, $p = .34$, $\eta_p^2 = .00$, CI [.00, .03]), and there was no interaction ($F(1, 207) = .98$, $p = .32$, $\eta_p^2 = .00$, CI [.00, .03]).

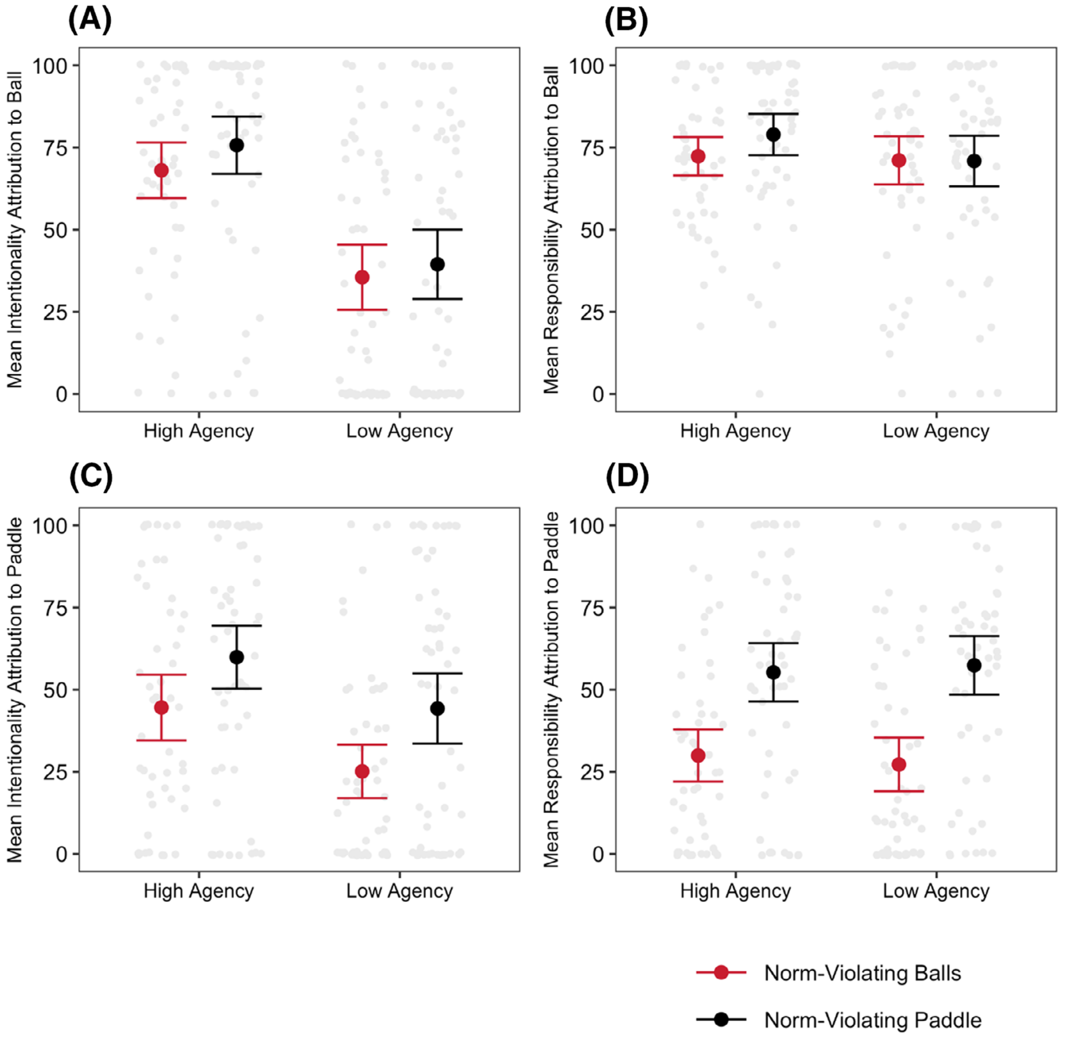

Fig. 6. Mean response for responsibility and intentionality questions in Experiment 4. (A) Mean intentionality attribution to ball A (low-agency conditions) or blue ball (high-agency conditions) in Experiment 4. (B) Mean responsibility attribution to ball A (low-agency conditions) or blue ball (high-agency conditions) in Experiment 4. (C) Mean intentionality attribution to paddle in Experiment 4. (D) Mean responsibility attribution to paddle in Experiment 4. Error bars indicate 95% confidence intervals. Light gray points represent individual participant responses evenly jittered.

5.2.3. *Intentionality attributed to paddle*

Participants in the high-agency conditions were more inclined to say that the paddle intentionally did not move to block the green ball from entering the tube than participants in the low-agency conditions were to say that the paddle intentionally did not move to block ball B from entering the tube ($F(1, 207) = 13.19$, $p < .001$, $\eta_p^2 = .06$, *CI* [.02, .12]) (Fig. 6C). Participants were also more inclined to attribute intentionality to the paddle's lack of movement in the norm-violating paddle conditions ($F(1, 207) = 13.50$, $p < .001$, $\eta^2_p = .06$, *CI* [.02, .12]) (Fig. 6C). There was no interaction ($F(1, 207) = .15$, $p = .69$, $\eta^2_p = .00$, *CI* [.00, .02]).

5.2.4. Responsibility attributed to paddle

We had no evidence that participants attribute more responsibility to the paddle as a function of the agency manipulation ($F(1, 207) = .00$, $p = .95$, $\eta^2_p = .00$, *CI* [.00, .00]). Participants were, however, more inclined to attribute responsibility to the paddle in the norm-violating-paddle conditions, relative to the norm-violating-balls conditions ($F(1, 207) = 42.77$, $p < .001$, $\eta^2_p = .17$, *CI* [.10, .25]) (Fig. 6D). There was no interaction ($F(1, 207) = .32$, $p = .56$, $\eta^2_p = .00$, *CI* [.00, .02]).

5.2.5. Causal selection

Next, we investigated participants' causal-selection behavior as a function of the norm and agency manipulation. Consistent with the previous experiments, there was a main effect of the norm conditions such that participants selected the norm-violating factor as the cause more frequently than they selected the factor that did not violate the norm ($b = 3.85$, $SE = 1.05$, $z = 3.66$, $p < .001$, CI [1.78, 5.90]) (Fig. 7). Critically, there was no main effect of agency ($b = 1.14$, $SE = 1.17$, $z = .97$, $p = .33$, CI [−1.15, 3.44]) or an interaction between the norm and agency manipulation ($b = -2.10$, $SE = 1.25$, $z = -1.68$, $p = .09$, CI [−4.54, .34]).

5.2.6. Causal selection and intentionality

We then investigated the fit of a series of multinomial logistic regressions when adding the measures of intentionality and responsibility to the models. Adding the intentionality attributed to the ball to the model did not improve the fit of the model ($LRT = 20.64$, $p = .05$, $\Delta AIC = 3.32$). Adding the intentionality attributed to the paddle to the model significantly improved the fit ($LRT = 23.33$, $p = .02$, $\Delta AIC = .66$). In this model, the abnormal-selection effect was still significant ($b = 2.30$, $SE = 1.42$, $z = 1.62$, $p < .001$, CI [−.48, 5.08]).

5.2.7. Causal selection and responsibility

Adding the responsibility attributed to the ball to the model significantly improved the fit ($LRT = 32.23$, $p = .001$, $\Delta AIC = -8.23$). In this model, the abnormal-selection effect was no longer significant ($b = 5.00$, $SE = 3.85$, $z = 1.30$, $p = .19$, CI [−2.54, 12.53]). Adding the responsibility attributed to the paddle to the model significantly improved the fit ($LRT = 27.76$, $p = .005$, $\Delta AIC = 3.76$). In this model, the abnormal-selection effect was still significant ($b = 5.27$, $SE = 3.23$, $z = 1.63$, $p < .001$, CI [−1.04, 11.59]).

5.3. Discussion

In Experiment 4, we find abnormal-selection effects with statistical norms for prospective causal judgments. The high agency condition increased the perceived intentionality

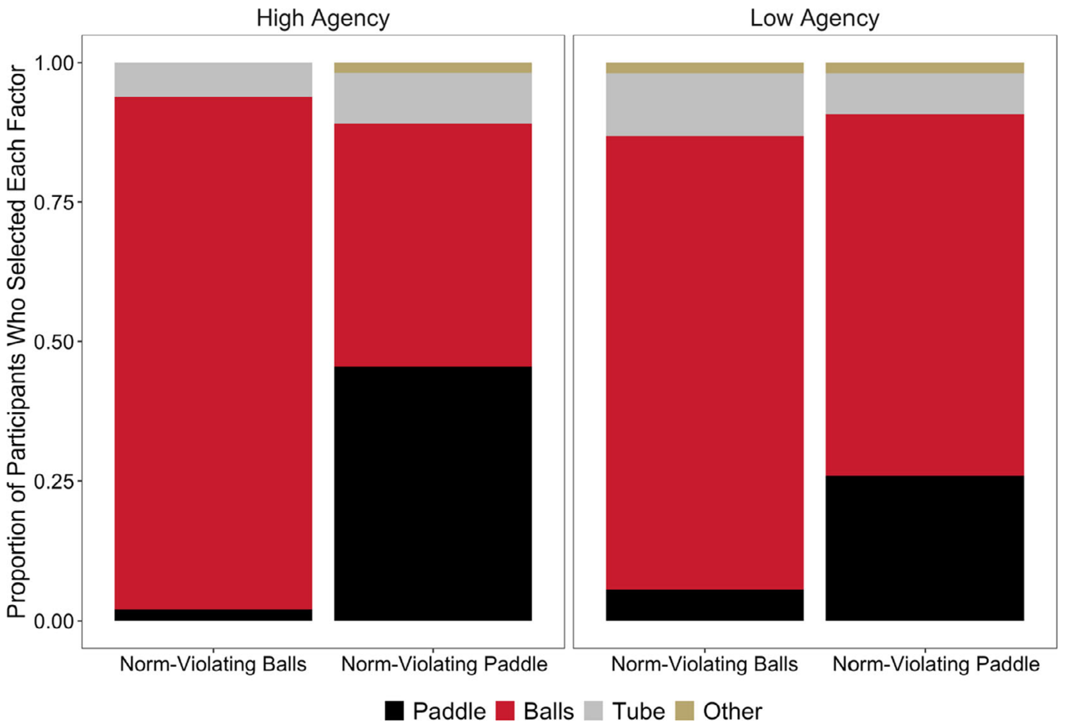

Fig. 7. Proportion of participants selecting each factor as the cause as a function of the norm and agency manipulation in Experiment 4.

of the individual events. The experiment, however, did not reveal a difference in participants' causal-selection behavior as a result of this manipulation; there is no evidence that abnormal-selection effects vary between high- and low-agency conditions as a function of perceived agency. While this evidence does not rule out the possibility that agency affects abnormal-selection behavior in purely agential scenarios or other contexts, it is strong evidence to support the modal explanation of abnormal-selection effects.

We do find that adding to the model participants' responsibility attributions to the ball slightly improves the model fit. This finding could suggest that attributions of responsibility have some role to play in causal-selection behavior (Livengood, Sytsma, & Rose, 2017; Sytsma, 2019, 2020; Sytsma, & Livengood, 2019; Sytsma, Livengood, & Rose, 2012). It could also be that participants' causal judgments affect their judgments of responsibility (see Henne et al., 2017). Likewise, we also see an increase in responsibility attribution for the paddle when it violates a norm (Fig. 6D). Again, this finding could be because an increase in causal selection would increase the degree of responsibility attribution, or it could be the other way around. We have no evidence, nonetheless, that these attributions of responsibility are agential because the responsibility attributions did not vary as a function of agency (Fig. 6B). And participants' understanding of "responsibility" may be ambiguous in the causal judgment literature between a causal-responsibility reading and an agentive-responsibility reading. Future work should determine the extent to which potential distinct notions of causal responsibility and agential responsibility and blame come apart.

Unexpectedly, we also found that there was an increase in intentionality attributed to the paddle's lack of movement when this lack of movement violated a statistical norm (Fig. 6C). The increase in abnormal-selection effects for prospective causal judgments could increase the extent to which participants see the omission as intentional. It could also be the other way around; an increase in perceived intentionality could increase the extent to which participants see the inaction as causal. While some researchers have explored whether norms affect perceived intentionality of side effects (Knobe, 2004; Uttich & Lombrozo, 2010), this effect, to our knowledge, has never been found for the inactions of non-agents. Future work should explore this unique finding.

6. General discussion

In four experiments, participants selected norm-violating factors more often than other factors as the prospective cause (Experiments 1–4). In Experiments 1 and 2, participants responded differently to some stimuli; the abnormal-selection effects were noticeably weaker for certain stimuli. We manipulated the temporal order of the factors in Experiment 3 to determine whether a recency effect explained these unexpected differences. The results of Experiment 3 show an overall abnormal-selection effect and a recency effect, suggesting that recency may explain the unexpected pattern of results from some items in Experiments 1 and 2. Experiment 4 shows that even when perceived agency affects judgments of intentionality and responsibility, we see consistent abnormal-selection behavior: People select norm-violating factors as the prospective cause at the same rate in high- and low-agency conditions.

The results of these experiments have some important consequences for the study of causal cognition. While accounting for some of the limitations of past work on abnormal selection, we present strong evidence in support of modal explanations for abnormal-selection effects. Participants in our studies select norm-violating factors as causes for stimuli that reduce the presence of agential cues (Experiments 1–3), and increasing agency cues does not change this tendency (Experiment 4). Social explanations might account for abnormal-selection behavior in some contexts, but, in general, abnormal-selection behavior likely does not depend on perceived intentions of agents, assessments of blame, or other social concerns. Rather, abnormal-selection effects seem to reflect a more general causal reasoning process, not just processes related to social or moral cognition that involves modal cognition. The modal explanations for abnormal-selection effects predict the results that we present here; in nonsocial situations, abnormal-selection effects should occur, and they should occur for prospective causal judgments. Even if the social explanation can account for the results of Experiments 1–3, it does not predict the results of Experiment 4. In Experiment 4, we increased agency cues, and we saw an increase in perceived intentionality attributed to the objects in our stimuli. But we did not see a change in abnormal-selection behavior, as social explanations predict. While these results are not evidence that the social explanation is completely mistaken about causal-selection behavior, we have strong evidence that modal explanations account for these effects—even when agency cues are increased.

Most critically, these results offer strong support for modal explanations of causal selection. This study is the first of its kind to investigate abnormal selection for prospective causal judgments. Most studies that investigate abnormal-selection effects—and causal judgments more generally—have examined retrospective causal judgments (Henne et al., 2017; Hitchcock & Knobe, 2009; see also Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017). In most studies, researchers give participants a scenario, and then they ask them what "caused" the outcome in the past. In these cases, there is a fact of the matter; for instance, both cars went through the intersection and an accident occurred. So, when people consider possible alternatives, they are relative to an actual situation that occurred. In our study, we asked participants about the cause of a future outcome (i.e., about what factor "will cause the ball to enter the goal"). While the potential causal factors occurred in the scenarios, the outcome did not. Analogously, the two cars went through the intersection, but the crash had yet to occur. In such cases, there is no fact of the matter to which participants consider an alternative. Yet our results still show patterns of abnormal-selection effects. As mentioned in Section 1, it seems challenging for social-cognition explanations to account for these findings because the outcome has yet to occur. We accept that some of these views may be readily modified to account for anticipated blame (Alicke et al., 2011; Rose, 2017). But the burden of explaining the mechanism is on those who accept social-cognition explanations. Modal explanations, in contrast, readily explain how norms affect prospective causal judgments, not just retrospective ones.

There are some competing modal explanations that readily account for the effect of norms on prospective causal reasoning. One recent counterfactual model, which we will refer to as the *necessity-sufficiency model*, explains the abnormal-selection effects for retrospective causal judgments (Icard et al., 2017). The core of this computational model is that norms affect people's counterfactual reasoning, thereby impacting the degree to which people focus on the necessity or sufficiency of the potential causal factor. On this view, all causal judgment involves thinking both about whether a factor is necessary for the outcome and about whether that factor is sufficient for the outcome. The effect of norms is then explained in terms of the comparative weight of necessity and sufficiency. The more abnormal people perceive a factor to be, the more people focus on the necessity of the abnormal factor. The more normal people perceive a factor to be, the more people focus on the sufficiency of the normal factor. The necessity-sufficiency model could account for our findings for prospective causal judgments. After the paddle moves, participants might consider the possibility, "what if the paddle had not moved," and ask if the ball would enter the goal in this hypothetical. So, prospective hypotheticals like these show that some factors in our stimuli are necessary for the ball to enter the goal in the future. On the necessity-sufficiency model, the abnormal factors in our experiments are seen as *more* necessary, thus *more* causal. Thus, we see abnormal-selection effects. Future work will have to see if such an adjustment to this model is plausible.

A model that accounts for the degree of necessity and sufficiency of a factor could explain some of the differences in judgments across our materials. In Experiment 2, participants were generally more inclined to say that the paddle's movement in V2 and V4 will cause the ball to enter the goal than they were to say that the paddle's failure to move in V1 and V3 will cause the ball to enter the goal. While the paddle's movement in V2 and V4 and the paddle's failure to move in V1 and V3 are both necessary for the ball to enter the goal, there is a difference in the degree of sufficiency; if the paddle moves in V2 and V4, the ball will enter the goal, but if the paddle fails to move in V1 and V3, the ball will not definitely enter the goal, as the tube has two outputs and only one faces the goal. Hence, there is a difference in the degree to which the particular factors across items are sufficient, engendering a difference in the overall pattern of causal judgments. The extent to which this difference is explained by a difference in sufficiency, rather than an action-inaction difference, will have to be explored in future work.

Another modal account of causal reasoning has the machinery to explain both retrospective and prospective causal cognition. It posits that people reason about causal relations by constructing a set of temporally ordered simulations of possibilities—that is, a set of "models" (Johnson-Laird, Khemlani, & Goodwin, 2015; Khemlani, Barbey, & Johnson-Laird, 2014). When thinking retrospectively about one event having caused an outcome, one of the models in the set of possibilities corresponds to a *fact*, and, therefore, all the other models in the set correspond to counterfactual alternatives (Khemlani, Byrne, & Johnson-Laird, 2018). When thinking prospectively about the potential for an event to cause an outcome in the future, the possibilities in the set serve as hypothetical —not counterfactual—alternatives. Hence, the model theory provides a uniform account for reasoning about past and future causes (Johnson-Laird & Khemlani, 2017), and it accords with recent theoretical frameworks that suggest modal cognition is central to reasoning about causality and morality (Phillips, Morris, & Cushman, 2019). Future work will have to compare the predictions of these distinct modal accounts.

7. Conclusion

In four experiments, we find an abnormal-selection effect for prospective causal judgment: People are more inclined to select norm-violating factor as the cause of some future outcome. While these studies account for many limitations of past work on abnormal-selection effects and provide evidence that these effects are robust, they also open up a range of questions for modal explanations of abnormal-selection effects. Future work will have to determine which of these modal accounts can explain these new results in conjunction with the large body of work on retrospective causal judgments.

Open Research badges

This article has earned Open Data and Open Materials badges. Data and materials are available at https://osf.io/4pvyd/.

Notes

1. Alicke and colleagues' model, for instance, "assumes that evaluative reactions accompany virtually all human events in which good or bad actions or outcomes occur" (Alicke, Rose, & Bloom, 2011). But if the outcome has yet to occur, then there has yet to be an evaluative reaction that motivates particular causal judgments. We acknowledge below that the culpable causation model may be amended to accommodate anticipated blame or reactions to future outcomes, but the model as it is formulated does not predict abnormal-selection effects for prospective causal judgments.
2. In Experiments 1–3, we deviated from the planned analyses in the preregistration. Our original, planned mixed-effects analysis did not allow us to use multinomial models—only binomial models. Because we required multinomial models for the investigation of our hypotheses, we deviated from our preregistered plan and used non-mixed effects multinomial models.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, *63*(3), 368.

Alicke, M. D., Mandel, D. R., Hilton, D. J., Gerstenberg, T., & Lagnado, D. A. (2015). Causal conceptions in social explanation and moral evaluation: A historical tour. *Perspectives on Psychological Science*, *10* (6), 790–812.

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

Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition*, *167*, 25–37.

Bello, P. F. (2016). Machine ethics and modal psychology. In V. C. Müller (Ed.), *Computing and philosophy* (pp. 245–258). Cham: Springer.

Bello, P., & Khemlani, S. (2015). A model-based theory of omissive causation. In R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 214–219). Austin, TX: Cognitive Science Society.

Bernstein, S. (2014). Omissions as possibilities. *Philosophical Studies*, *167*(1), 1–23.

Bernstein, S. (2015). The metaphysics of omissions. *Philosophy Compass*, *10*(3), 208–218.

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., & De Rosario, M. H. (2018). Package 'pwr'. R package version. 1–2.

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

Gerstenberg, T., & Icard, T. F. (2019). Expectations affect physical causation judgments. *Journal of Experimental Psychology: General*,*149*(3), 599–607.

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *28*(12), 1731–1744.

Halpern, J. Y., & Hitchcock, C. (2014). Graded causation and defaults. *The British Journal for the Philosophy of Science*, *66*(2), 413–457.

Hart, H. L. A., & Honoré, T. (1985). *Causation in the law*. Oxford: OUP Oxford.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243–259.

Henne, P., Bello, P., Khemlani, S., & De Brigard, F. (2019). Norms and the meaning of omissive enabling conditions. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the Cognitive Science Society* (pp. 1901–1907). Montreal, Canada: Cognitive Science Society.

Henne, P., Niemi, L., Pinillos, Á., De Brigard, F., & Knobe, J. (2019). A counterfactual explanation for the action effect in causal judgment. *Cognition*, *190*, 157–164.

Henne, P., Pinillos, Á., & De Brigard, F. (2017). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, *95*(2), 270–283.

Hesslow, G. (1988). The problem of causal selection. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 11–32). New York: New York University Press.

Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, *93*(1), 75.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy*, *106*(11), 587–612.

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition*, *161*, 80–93.

Johnson-Laird, P. N., & Khemlani, S. (2017). Mental models and causation. In M. Waldmann (Ed.), *Oxford handbook of causal reasoning* (pp. 1–42). New York: Elsevier Inc, Academic Press.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Response to Baratgin et al: Mental models integrate probability and deduction. *Trends in Cognitive Sciences*, *19*(10), 548–549.

Khemlani, S. S., Barbey, A. K., & Johnson-Laird, P. N. (2014). Causal reasoning with mental models. *Frontiers in Human Neuroscience*, *8*, 849.

Khemlani, S. S., Byrne, R. M., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, *42*(6), 1887–1924.

Khemlani, S., Wasylyshyn, C., Briggs, G., & Bello, P. (2018). Mental models and omissive causation. *Memory & Cognition*, *46*(8), 1344–1359. https://doi.org/10.3758/s13421-018-0841-5

Kirfel, L., & Lagnado, D. (2018). Statistical norm effects in causal cognition. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.). *Proceedings of the 40th annual conference of the Cognitive Science Society* (pp. 615–620). Madison, WI: Cognitive Science Society.

Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis*, *64*(2), 181–187.

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

Kominsky, J. F., & Phillips, J. (2019). Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science*, *43*, e12792. https://doi.org/10.1111/cogs.12792

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

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*(3), 754–770.

Lewis, D. (1974). Causation. *The Journal of Philosophy*, *70*(17), 556–567.

Livengood, J., Sytsma, J., & Rose, D. (2017). Following the FAD: Folk attributions and theories of actual causation. *Review of Philosophy and Psychology*, *8*(2), 273–294.

McGrath, S. (2005). Causation by omission: A dilemma. *Philosophical Studies*, *123*(1), 125–148.

Morris, A., Phillips, J., Gerstenberg, T., & Cushman, F. (2019). Quantitative causal selection patterns in token causation. *PLoS One*, *14*(8), e0219704.

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.

Phillips, J., Morris, A., & Cushman, F. (2019). How we know what not to think. *Trends in Cognitive Sciences*, *23*(12), 1026–1040.

Reuter, K., Kirfel, L., Van Riel, R., & Barlassina, L. (2014). The good, the bad, and the timely: How temporal order and moral judgment influence causal selection. *Frontiers in Psychology*, *5*, 1336.

Rogers, R., Alicke, M. D., Taylor, S. G., Rose, D., Davis, T. L., & Bloom, D. (2019). Causal deviance and the ascription of intent and blame. *Philosophical Psychology*, *32*(3), 404–427.

Rose, D. (2017). Folk intuitions of actual causation: a two-pronged debunking explanation. *Philosophical Studies*, *174*(5), 1323–1361.

Schaffer, J. (2013). Causal contextualisms. In M. Blaauw (Ed.), *Contrastivism in philosophy* (pp. 43–71). New York: Routledge.

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York, NY: Springer-Verlag.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General*, *126*(4), 323–348.

Sytsma, J. (2019). Structure and norms: Investigating the pattern of effects for causal attributions. [Preprint]. http://philsci-archive.pitt.edu/16626/

Sytsma, J. (2020). Causation, responsibility, and typicality. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-020-00498-2

Sytsma, J., & Livengood, J. (2019). Causal attributions and the trolley problem. Preprint. http://philsci-archive.pitt.edu/16200/

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.

Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition*, *116*(1), 87–100.