Norms Affect Prospective Causal Judgments

Paul Henne*

Department of Philosophy, Lake Forest College

Kevin O'Neill

Center for Cognitive Neuroscience, Duke University

Paul Bello

Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory

Sangeet Khemlani

Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory

Felipe De Brigard

Department of Philosophy, Duke University

Department of Psychology and Neuroscience, Duke University

Center for Cognitive Neuroscience, Duke University

Duke Institute for Brain Sciences, Duke University

*Correspondence to:

Paul Henne
Durand Art Institute
Lake Forest College
555 North Sheridan Road
Lake Forest, IL 60045
phenne@lakeforest.edu

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 only retrospective vignette-based paradigms. In within-participants designs, we use a novel set of videos to investigate this effect for prospective causal judgments—i.e., judgments about the cause of some future outcome. Three experiments show that people more frequently select norm-violating factors, relative to norm-conforming ones, as the cause of some future outcome. 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. This difficulty in describing why people select a particular factor as the cause over all the other necessary factors is called *the selection problem* (Hesslow, 1988; Hart & Honore, 1985).

Until recently, theorists dismissed the selection problem as an issue of pragmatics, irrelevant to understanding the meaning of causal statements. Philosopher 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: Schaffer, 2013; Driver, 2008). However, a number of cognitive scientists and philosophers have recently turned their attention to the selection problem and have sought to explain it (Bernstein, 2014; 2015; Henne, Pinillos, & De Brigard, 2017; Hitchcock & Knobe, 2009; McGrath, 2005).

Some of the researchers who attempt to explain the selection problem argue that norms affect causal selection such that people select norm-violating factors as causes (following Hart & Honore, 1985). For these researchers, norms generally 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, so people select it relative to the other norm-conforming factors. And extensive empirical work shows that, in fact, people do tend to select norm-violating factors rather than norm-conforming factors as the cause (Gerstenberg & Icard, 2019; Hilton & Slugoski, 1986; Hitchcock & Knobe, 2009; Knobe & Fraser, 2008), both in the case of actions (Phillips, Luguri, & Knobe, 2015) and in the case of omissions (Henne, Pinillos, & De Brigard, 2017). Although these findings are consistent, they differ when participants are asked about alternative potential causal factors—i.e., necessary factors that are fixed rather than varied as norm-violating or norm-conforming—and when the varied factors are more or less necessary or sufficient for the outcome. When multiple factors are jointly sufficient for the outcome, for instance, the presence of norm-violating factors decreases the extent to which people agree that alternative potential causal factors caused the outcome (Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Morris, Phillips, Gerstenberg, & Cushman, 2019). When factors are individually sufficient but not individually necessary for the outcome, moreover, people tend to select norm-conforming factors as the cause (Icard, Kominsky, & Knobe, 2017; Kirfel & Lagnado, 2018; Morris et al., 2019). Recent computational and formal frameworks of causal reasoning represent norms directly to account for these effects (Bello, 2015; Halpern & Hitchcock, 2015; Icard, Kominsky, & Knobe, 2017), and some related frameworks have predicted and explained other features of causal judgment such as action-inaction differences (Henne, Niemi, Pinillos, De Brigard, & Knobe, 2019).

Researchers have proposed several explanations for why people select norm-violating factors as causes. Some theorists argue that these abnormal-selection effects result from the inclination to blame agents (Alicke, 1992; Alicke, Rose, & Bloom, 2011; Alicke et al., 2015; Rogers et al., 2019) or to hold them responsible (Sytsma, Livengood, & Rose 2012; Livengood, Sytsma, & Rose 2017; Sytsma, & Livengood, 2019). On some such views, when making causal judgments people engage in blame validation: they align their evaluation of a causal relation with their desire to blame the source of their reactions (Alicke, 1992; Rogers et al., 2019). In the introductory example, these theories predict that people would judge that the car that went through the red light caused the crash because people align their causal attribution with their desire to blame the driver of the car that went through the red light. As such, blame or related features of social cognition affect causal judgments. 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, Rose, & Bloom, 2011; Alicke et al., 2015). Hence, abnormal-selection effects should not arise in non-social situations. While these views vary in a number of ways, we will generally refer to them as *social-cognition* explanations.

Other researchers argue that abnormal-selection effects reflect more general cognitive processes than social cognition (Hitchcock & Knobe, 2009; Icard, Kominsky, & Knobe, 2017; Henne, Bello, Khemlani, & De Brigard, 2019; Gerstenberg & Icard, 2019). These researchers argue that norms affect people’s explicit (Phillips, Luguri, & Knobe, 2015; Kominsky & Phillips, 2019) or implicit (Icard, Kominsky, & Knobe, 2017) consideration of relevant counterfactual alternatives (see also Morris et al., 2019), and because counterfactual thinking 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’ve happened), so they are more inclined to agree that the car going through the red light caused the crash. Some others in this group argue that while consideration of particular hypothetical alternatives determine causal judgments (Johnson-Laird & Khemlani, 2017; Khemlani, Wasylyshyn, Briggs, & Bello, 2018) norms may help people to select factors as candidate causes (Bello & Khemlani, 2015; Henne 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.

Are abnormal-selection effects best explained by social cognition, or are they a more general phenomenon related to reasoning about possibilities? In this article, our goal is to test two predictions of the modal explanation. First, we wanted to know if norms affect causal judgments in non-social, visual stimuli—that is, videos of physical objects interacting. The modal explanation predicts that they would, while the social-cognition explanation does not, as no blame or praise attributions are relevant in non-social scenarios. Second, we sought to test whether norms affect prospective causal judgments—i.e., judgments about the cause of some future outcome after the potential causal factors occur. That is, participants in our experiments did not see the outcome occur, and they selected whether some factor (event or omission) *will cause* the outcome. Social-cognition explanations hold that people align their evaluation of a causal relation with their desire to blame or praise the source of their reactions. That is, such explanations rely on the judgments being retrospective (Alicke, Rose & Bloom, 2014). So, if people engage in blame validation or excuse validation in making their causal judgments, then prospective causal judgments—particularly those where the outcome will not definitively occur—should not be affected by norms. Some of the modal explanations, however, predict that norms should affect prospective causal judgments just as they do for retrospective causal judgments. 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 counterfactuals 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.

A secondary objective of the current study is to overcome limitations of past work on abnormal selection. Previous work on abnormal-selection effects has three primary limitations. First, most studies that explored abnormal-selection effects have used verbal vignettes to prompt causal judgments (e.g., Henne, Pinillos, & De Brigard, 2017; but see: Gerstenberg & Icard, 2019; Henne et al., 2019; Kirfel & Lagnado, 2019). Vignette-based studies require participants to imagine the interactions between different entities and base their responses on what they imagine. Hence, it is very challenging to develop mechanistic theories that can explain the time course of causal reasoning from vignette-based studies. To deal with this first limitation, we used videos to investigate abnormal-selection effects. This change also allows us to control more tightly the independent variables in the experiments. Second, vignette-based studies often involve agents, which allows for potential interference of perceived intentionality and blame attribution. Such interference can conflate causal cognition with moral cognition or some other process altogether. To deal with the second limitation, we used videos without any agents. Instead, we employed balls colliding and a paddle to minimize participants' attributions of blame or intentionality to certain factors rather than others. Third, most studies investigating abnormal-selection effects use only a single vignette (e.g., Henne, Pinillos, & De Brigard, 2017; cf., Icard, Kominsky, & Knobe, 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. To deal with this last limitation, we used multiple videos with different configurations to test our predictions.

Accounting for the limitations of these previous experiments, we aimed to answer one question: do people select norm-violating factors as *the* cause of some outcome more frequently than they select norm-conforming ones in non-social situations across many items? In the three experiments below, we found consistent abnormal-selection effects. In Experiment 1, we found that norms affect causal selection; for multiple videos, participants are more inclined to select norm-violating factors, relative to norm-conforming ones, as the cause of the outcome. Experiment 2 sought to replicate the pattern of results from Experiment 1, extend it to novel stimuli, and rule out potential confounds. In Experiment 2, we again found that participants overall were more likely to select the norm-violating factors, relative to the norm-conforming ones, as the cause of the outcome. While there was an overall abnormal-selection effect, however, participants responded differently to some of the videos. Realizing that this different pattern of responses to some of the videos may be a result of a recency effect, we manipulated the temporal order of the items that elicited this different pattern of responses in Experiment 3. The results of Experiment 3 show an overall abnormal-selection effect and a recency effect. This manipulation explained why participants responded differently to some videos in Experiment 1 and Experiment 2.

These results have important consequences for the study of causal reasoning. First, people select norm-violating factors as causes even for stimuli that reduce the presence of agential cues. Hence, the effects reflect general causal reasoning processes, not just processes related to social or moral cognition. Second, participants in our studies select norm-violating factors in real-time, and this novel methodology and stimuli allow us to track real-time causal selection behavior. Future studies will examine the time course of online causal selection, which may inform the development of more detailed mechanistic accounts. Third, this study is the first of its kind to investigate abnormal selection for prospective causal judgments. Any modal explanation of abnormal selection will have to account for these new findings.

2. Experiment 1

Experiment 1 tests the two predictions of the modal explanation for abnormal-selection effects. First, it examines if prescriptive norms affect participant's selection of causes in visual stimuli that are non-social. Specifically, it investigates whether participants are more likely to select norm-violating factors than they are to select those factors where no norm is specified. Second, it examines whether abnormal-selection effects arise for prospective causal judgments.

To these ends, we have participants watch three videos of different object configurations in which two balls (A and B) collide and where—depending on the video—a paddle moves to block a tube or does not move to block a tube. For each of the three videos, the balls colliding or the paddle moving (or not moving) are norm-violating. That is, participants are told that the collision of the balls or the paddle's movement (or failure to move) is not supposed to occur. Participants watch all of the videos, thus they see all object configurations and all norm manipulations. Each video ends before ball A enters a goal, and participants are asked to select from a list of options which factor *will* cause ball A to enter the goal. The modal explanation predicts that norm-violating factors should be selected as the cause more often than norm-conforming factors. All materials, preregistrations, and analysis code for all experiments are available at https://osf.io/4pvyd/.

2.1 Methods

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 3 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 = .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 are located in the United States and have not participated in a pilot experiment. We excluded 10 participants for not paying attention and analyzed the data from 275 remaining participants (M_{age} = 35.30, $\text{Range}_{\text{age}}$ = [19-74], 35.60% female).

Materials and Procedure. We assigned participants to all 6 study conditions in randomized order in a 3 (Item: V1, V2, V3) x 2 (Norm: Balls or Paddle) within-participants design (Figure 1). Using still images of V1 (video 1), the study familiarized participants with the videos and with the 4 components of the videos: the balls, the paddle, the goal, and the tube. The study further familiarized participants with the components by watching videos with the object orientation of V1. All participants watched three videos of the V1 arrangement and read a description of what occurred in the video. Specifically, they 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. The study then told participants that the instruction phase was over and that the test phase would start. Before each video, the study told participants that either the balls colliding was norm-violating (e.g., "In the video you're about to watch the balls are **not** supposed to collide.") or the paddle not moving (or moving as in V2) was norm-violating (e.g., "In the video you're about to watch the paddle is **not** supposed to block the tube."). To verify that participants understood what the norm was, the study directly asked participants 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?"). Participants answered "Yes" or "No." If they answered correctly, they received confirmatory feedback and then were reminded of the norm. If they answered incorrectly, they received corrective feedback and then were reminded of the norm. On the following page, the study 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 then participants watched one video. Below the video, they read, "**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 factor from a drop-down menu that listed the following factors: the balls colliding, the paddle moving (or not moving), the tube, or some other event. The wording for each factor varied slightly for each video. For instance, the options for V1 were: "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 study asked participants for basic demographic information. The study then asked participants to respond to an explicit attention check.

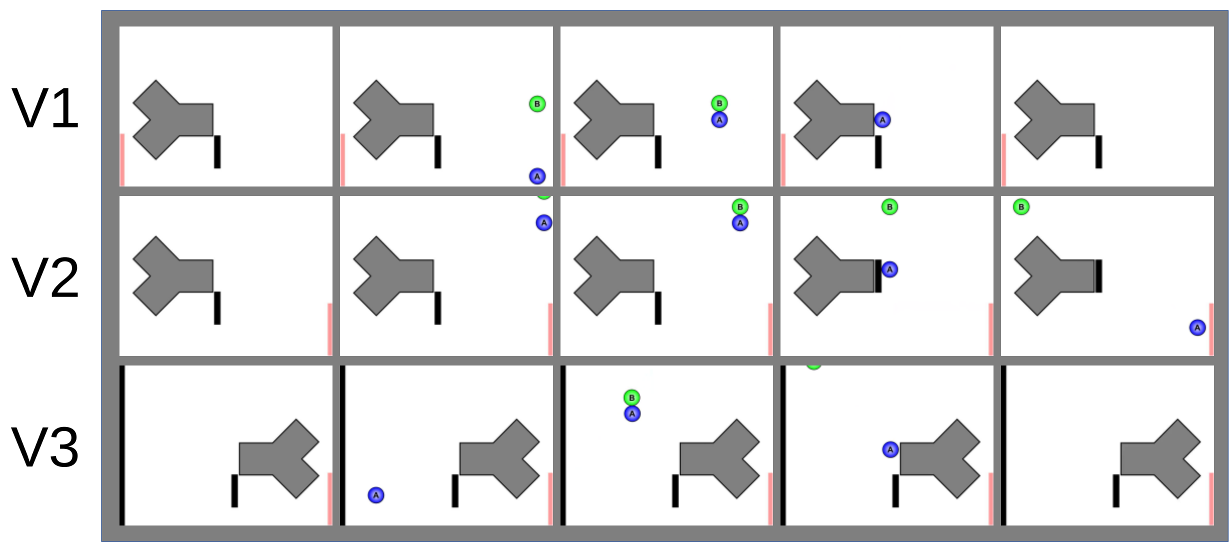

Figure 1. The three videos used in Experiment 1. In each video, a blue ball (A) collides with a green ball (B) and moves in the direction of a grey forked tube. The tube can be blocked by the black paddle. The pink bar represents the goal.

2.2 Results

Data collection was completed for each experiment in this manuscript prior to any analysis of the data for each experiment. Data were fit to mixed-effects models (random slopes and intercepts) using the lme4 software package (Bates, Maechler, Bolker, & Walker, 2015).

To test whether participants were more likely to select norm-violating factors, we first examined the difference in the selection of the potential causal factors as a function of the norm condition across all three videos. Given that each potential causal factor varied across each item (V1, V2, V3), we coded all factors that participants selected as either ‘balls’, ‘paddle’, ‘tube’, or ‘other’. In line with norm-based accounts of causal selection, the norm conditions affected participants selection of the cause (*ß* = 2.05, *SE* = .18, *z* = 11.41, $p < .001$, CI [1.71, 2.42]) (Figure 2). We report the percentage of participants who selected each factor in each video in Table 1.

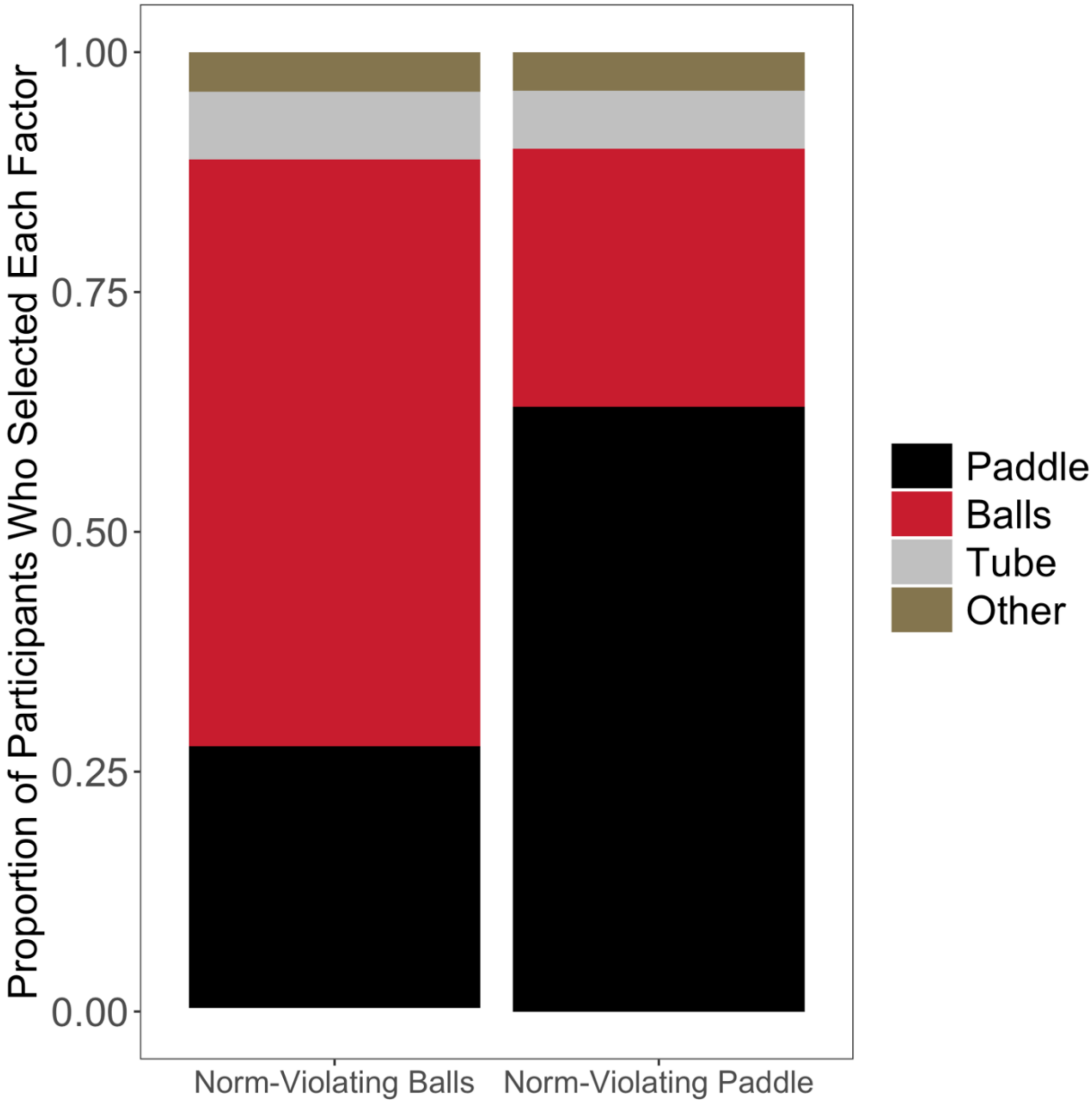

Figure 2. Proportion of participants who selected each factor as the cause as a function of condition collapsed across all videos in Experiment 1.

	V1		V2		V3	
	Norm-Violating Balls	Norm-Violating Paddle	Norm-Violating Balls	Norm-Violating Paddle	Norm-Violating Balls	Norm-Violating Paddle
Balls	77%	28.5%	31.1%	14.7%	76%	37.1%
Paddle	11.3%	59.9%	59.3%	75.8%	11.6%	53.8%
Tube	9.1%	10.2%	2.2%	.7%	9.8%	7.3%
Other	2.6%	1.5%	7.3%	8.8%	2.5%	1.8%

Table 1. Percentage of participants' selection of each factor as the cause in Experiment 1.

2.3 Discussion

In line with modal explanations of abnormal-selection effects, participants more frequently selected norm-violating factors than they were to select a factor that did not violate a norm. The pattern arose across multiple items with visual stimuli that are non-social and that do not include agents. This result supports the view that perceived norms helps people select between causes and causal conditions, and it is consistent with the modal explanation.

There appeared to be a difference in participants' response between one item (V2) and the others (V1 and V3). For V2, participants seemed to be much more inclined to select the paddle's movement—rather than the balls colliding, the tube, or something else altogether—as the cause—even when the balls colliding is norm-violating. To investigate this apparent difference, we performed an unplanned analysis. We removed Item as random intercept in the model, and then we looked for an interaction between norms and items. While there was a significant interaction between norms and items, a post-hoc pairwise comparison revealed that participants were not significantly more inclined to select norm-violating factors for V2 ($ß = -.82$, $SE = .29$, $z = -2.77$, $p = .06$, CI [-1.66, .02]). This result was unexpected. Because we randomly selected the videos and their unique orientations of the balls colliding and the paddle's movement, we were unsure why participants responded differently to V2.

One potential explanation for this unexpected result is that there may have been some confounds unrelated to our experimental task. First, we worried that participants only saw the paddle move two times in the experiment (once during the test 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 statistical normal. This perceived difference in statistical normality could have affected the selection of the cause (Gerstenberg & Icard, 2019). Second, we noticed that the orientation of the videos differed between items. This difference could have engendered the difference in responses. We eliminated these two confounds in Experiment 2.

3. Experiment 2

Experiment 2 served to replicate Experiment 1, i.e., it sought to determine if norms affect people's prospective causal judgments in non-social visual stimuli. Unlike Experiment 1, Experiment 2 controlled for the statistical normality of the paddle moving or not moving by adding a video (V4) where the paddle also moved. By adding this video, participants, thus, 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) collide and where a paddle moves to block a tube or does not move to block a tube. As in Experiment 1, all videos end before ball A enters a goal, and we asked participants to select from a list which factor will cause ball A to enter the goal. For each of the four videos, the experiment varied abnormality of the balls colliding or the paddle moving. We took one video directly from Experiment 1 (V3), we changed the orientation of two videos (V1 and V2), and we created one new video (V4).

3.1 Methods

Participants. We kept the same sample size as that in Experiment 1. So, 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 are located in the United States and have not participated in Experiment 1 or a pilot experiment. We excluded 21 participants for not paying attention and analyzed the data from 359 remaining participants (M_{age} = 32.80, $\text{Range}_{\text{age}}$ = [18-64], 46.50% female).

Materials and Procedure. We assigned participants to all 8 conditions in randomized order in a 4 (Item: V1, V2, V3, V4) x 2 (Norm: Balls or Paddle) within-participants design. The study familiarized each participant with the videos and the 4 components of the videos just as in Experiment 1. The study then told participants that the instruction phase was over and that the test phase would start. The experimental design was exactly the same as that in Experiment 1. The only difference was the additional, modified materials included. After participants completed this task for all eight videos, the study asked them for basic demographic information. The study then asked participants to respond to the same explicit attention check used in Experiment 1.

3.2 Results

Data were analyzed using the same software and regression formula as in Experiment 1. To test whether participants were more likely to select norm-violating factors, we first examined the difference in the selection of the potential causal factors as a function of the norm condition across all four videos. Given that each potential causal factor varied slightly across each item (V1, V2, V3, V4), we coded all factors that participants selected as either ‘balls’, ‘paddle’, ‘tube’, or ‘other’. Participants selected norm-violating factors as the cause more often than norm-conforming factors (*ß* = 2.5, *SE* = .17, *z* = 14.03, $p < .001$, CI [2.17, 2.87]), even for material that had no obvious social context. The results corroborate modal explanations but challenge social-cognition explanations (Figure 3). We report the percentage of participants who selected each factor in each video in Table 2.

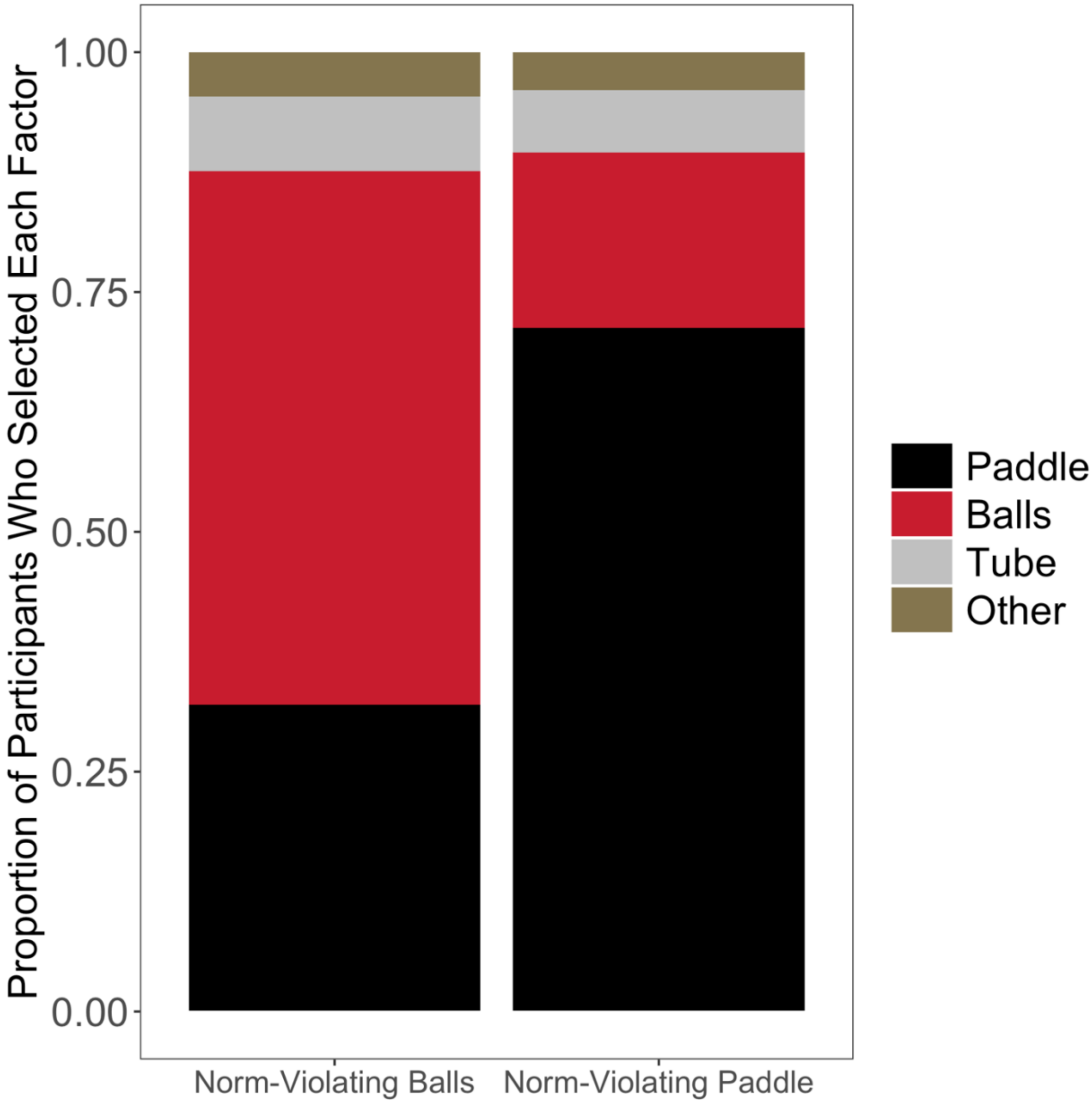

Figure 3. Proportion of participants selecting each factor as the cause as a function of condition collapsed across all videos in Experiment 2.

	V1		V2		V3		V4	
	Norm-Violating Balls	Norm-Violating Paddle	Norm-Violating Balls	Norm-Violating Paddle	Norm-Violating Balls	Norm-Violating Paddle	Norm-Violating Balls	Norm-Violating Paddle
Balls	70.5%	23.7%	41.3%	10.9%	71.6%	28.7%	39.4%	10.1%
Paddle	13.6%	62.4%	50.6%	80.7%	11.1%	60.4%	52.5%	81.3%
Tube	12.5%	11.4%	2.5%	2.2%	13.4%	8.6%	2.8%	3.6%
Other	3.3%	2.5%	5.6%	6.1%	3.9%	2.2%	5.3%	5%

Table 2. Percentage of participants' selection of each factor as the cause in Experiment 2.

3.3 Discussion

Again, our result showed that participants selected norm-violating factors as the cause more often than other factors. That is, abnormal-selection effects arose across multiple, non-social visual stimuli for prospective causal judgments. This new evidence supports the view that perceived norms helps people 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 orientation, there remained an apparent difference in participants' response between two items (V2 and V4) and the others (V1 and V3). 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. To further investigate this apparent difference, we performed the same post-hoc analyses as in Experiment 1. Tukey corrected, pairwise comparisons confirmed that participants were significantly more inclined to select norm-violating factors than other factors for all items (all $p < .0001$). Nonetheless, the proportion of participants selecting the paddle's movement is higher in V2 and V4. 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, & V4), it is still curious that the proportion of responses was different between V1 and V3 relative to V2 and V4.

One alternative explanation for this difference in selection across items is that a recency effect also affected participants' judgments in our materials. In V2 and V4, the paddle moves after the balls collide. In V1 and V3, the paddle does not move at all. Some recent work shows that people are more inclined to select the most recent factors as the cause of the outcome (e.g., Reuter, Kirfel, Van Riel, & Barlassina, 2014). 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 investigated this possibility.

4. Experiment 3

In our previous experiment, we found that participants were generally more inclined to select the paddle's movement as the cause of the outcome in some of our materials (V2 and V4). In these materials, the paddle moving was the most recent event before the outcome. Because people are more inclined to select recent events as causes (Byrne, 2016; Reuter et al., 2014), it might have been the case that a recency effect interfered with our investigation of norms. In this follow-up experiment, we test whether this different pattern occurred because of the temporal order of the factors.

In order 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, we made the same prediction about norms: if norms affect the selection of a potential causal factor, then participants will select the norm-violating factor more frequently than any others. Moreover, if temporal order affects the selection of a potential causal factor, then participants will select the more recently occurring factor more frequently than any others. Lastly, we predicted that, in situations where the paddle moves first and the balls' collision is norm-violating, participants would select the balls most frequently as the cause.

4.1 Methods

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 are located in the United States and have 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 for not paying attention and analyzed the data from 365 remaining participants (M_{age} = 31.5, $\text{Range}_{\text{age}}$ = [18-66], 50.8% female).

Materials and Procedure. We assigned participants to all 8 conditions in randomized order in a 2 (Item: V2 or V4) x 2 (Norm: Balls or Paddle) x 2 (Order: Balls First or Paddle First) within-participants design. The study familiarized each participant with the videos and the 4 components of the videos just as in Experiment 2. The study then told participants that the instruction phase was over and that the test phase would start. The experimental design was exactly the same as that in Experiment 2 in terms the norm manipulation. The only difference was that V2 and V4 from Experiment 2 were manipulated so that either the paddle moved first or the balls collided first. Hence, participants in this experiment watched the same videos from Experiment 2 (V2 and V4) and the new manipulated versions. After participants completed this task for all eight videos, they were asked for basic demographic information. They then responded to the same explicit attention check used in Experiment 2.

4.2 Results

We attempted to analyze the data using the same software and models as in Experiments 1 and 2. However, given the smaller number of distinct items, the mixed models converged. So, we simplified the models, and we used only Generalized Linear Models rather than mixed models. First, there was a main effect for norms; participants selected norm-violating factors as the cause more often than they selected other factors ($ß$ = .88, SE = .12, z = 7.37, p < .001, CI [.65, 1.12]).

There was also a main effect of order; people selected more recent factors as the cause (*ß* = -.54, *SE* = .1, *z* = -5.14, *p* < .001, CI [-.75, -.33]). There was no interaction of norms and order (*ß* = .03, *SE* = .16, *z* = .23, *p* = .81, CI [-.28, .35]). Lastly, we performed planned pairwise comparisons on all conditions. Notably, in the norm-violating balls conditions the collision of the balls was more frequently selected as the cause when the paddle moved first (Figure 4).

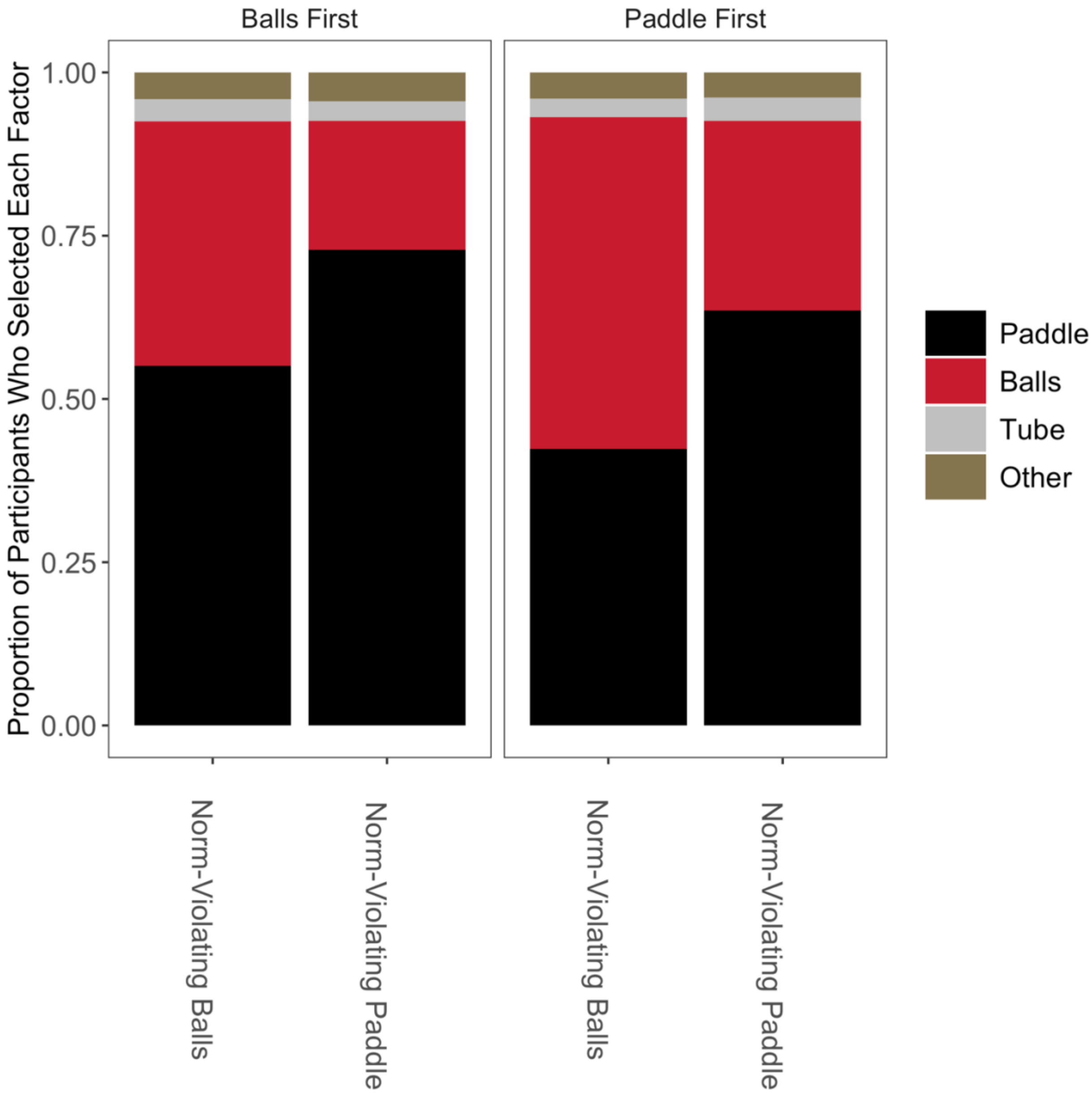

Figure 4. Proportion of participants selecting each factor as the cause as a function of condition collapsed across all videos in Experiment 3.

4.3 Discussion

Consistent with Experiments 1 and 2, participants in Experiment 3 selected norm-violating factors more often than they selected other factors. This response pattern again arose again across multiple items with visual stimuli that are non-social and that do not include agents. In Experiment 3, we also found a recency effect: participants were more inclined to select the most recent factor as the cause of the outcome. These finding suggests that recency at least partially explains the differences between items that we see in Experiment 1 and 2.

5. General Discussion

In three experiments, we find evidence that prescriptive norms resolve the selection problem. Participants selected norm-violating factors more often than other factors as the cause (Experiment 1, 2, and 3). In Experiment 1 and Experiment 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 if a recency effect explained these unexpected differences. The results of Experiment 3 show an overall abnormal-selection effect and a recency effect. This manipulation explained the unexpected pattern of results from some items in Experiment 1 and Experiment 2.

These results of these experiments have major consequences for the study of causal cognition. First, we have accounted for many of the limitations of previous studies on abnormal selection. Thus, we show that abnormal-selection effects are robust. Participants in our studies responded to questions about videos, rather than vignettes, which are typically used. We also used multiple items in order to be sure that these effects do not result from participants' responses to a single item. Hence, our experiments expand the generalizability of abnormal-selection effects on causal judgment.

Second, we present strong evidence against social-cognition explanations of abnormal-selection effects. Our results show that people select norm-violating factors as causes for stimuli that reduce the presence of agential cues. Hence, abnormal-selection effects reflect general causal reasoning processes, not just processes related to social or moral cognition, so abnormal-selection effects likely do not depend on perceived intentions of agents or other social properties. These findings for prospective causal judgments lend support for modal explanations of abnormal selection effects over social explanation. The modal explanations for abnormal selection effects predict the results that we present here; in non-social situations, abnormal-selection effects should occur, and they should occur for prospective causal judgments. The social explanation, however, does not predict the results of our experiments. In the non-social situations, no blame or praise is relevant to the causal judgment. Moreover, participants in our studies made the prospective causal judgment before the outcome occurred. For some of our materials (V1 and V3), participants do not know for sure that the ball will enter the goal. In such cases, there is no definitive outcome for participants to respond to and align their causal judgments with. Hence, there is no evidence that blame and excuse validation models can account for the abnormal-selection effects in our studies. While not all such blame or responsibility views make this prediction (see Sytsma, 2019), we have evidence that is generally consistent with the modal explanations for these effects.

Some researchers may want to push back on this second point. Specifically, some may worry that the stimuli employed in the current experiments did not eliminate agency cues entirely, thus a potential social-cognition explanation cannot be ruled out completely. After all, people perceive even basic geometric figures as agents engaging in social interactions (e.g., Heider and Simmel, 1944). Nevertheless, while we acknowledge the possibility that participants may interpret some of the factors in our stimuli as social or intentional agents, we do not believe that a social-cognition explanation would account for the findings. On the one hand, our stimuli differ in important respects from the kinds in Heider and Simmel's studies. While in the latter judgments of agency tend to emerge after watching the geometric figures move about for a while in unpredictable, intentional patterns, our stimuli are much shorter in duration and do not involve unpredictable, intentional movements, as they are inanimate physical objects bumping into one another in physically predictable ways. Thus, it is unlikely that participants saw the potential causal factors in our experiments as agents. On the other hand, if it is the case that even these basic causal stimuli are perceived as involving agency, then one worries that all causal scenarios could be perceived as social-agential interactions, in which case the social-agential explanation loses power to account for asymmetries in causal judgments. These considerations prevent us, as we say in the introduction, from claiming that the findings in the current experiments provide equal evidentiary value against the social-congition explanation as they do in favor of the modal explanation. Future studies where agency is definitively controlled for or directly manipulated will be needed to settle this difficult issue.

Third, these experiments allow researchers to build mechanistic accounts of causal selection. Participants in our experiments select norm-violating factors in real-time; the abnormal factors occur, and participants select them as the cause of the future outcome (i.e., entering the goal). This novel design allows us to investigate real-time causal selection behavior. Future studies can examine the time course of online causal selection, which can inform the development of more mechanistic accounts.

Fourth, the results of these experiments have consequences for the modal accounts 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 investigated retrospective causal judgments (Henne, Pinillos, De Brigard, 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 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"), and our results still show patterns of abnormal-selection effects. Modal explanations must account for how norms affect prospective causal judgments, not just retrospective ones.

There are some competing modal explanations that can potentially 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, Kominsky, & Knobe, 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 be adjusted to account for our findings for prospective causal judgments. After the paddle moves, participants might consider the counterfactual, "what if the paddle had not moved," and ask if the ball would enter the goal in this counterfactual. So, prospective counterfactuals 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.

One 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, 2015). 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.

6. Conclusion

In three within-participants 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.

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.

Bates, D., Maechler, M., & Bolker, B. S. Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

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

Bello, P. F. (2016). Machine ethics and modal psychology. In *Computing and philosophy* (pp. 245-258). Springer, Cham.

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*. 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, US: MIT Press.

Gerstenberg T., Icard T. F. (2019). Expectations affect physical causation judgments. In *Journal of Experimental Psychology: General*.

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*. OUP Oxford.

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*. 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. *Contemporary science and natural explanation: Commonsense conceptions of causality*, 11-32.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American journal of psychology*, *57*(2), 243-259.

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.

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

Johnson-Laird, P.N., & Khemlani, S. (2017). Mental models and causation. In M. Waldmann (Ed.), *Oxford Handbook of Causal Reasoning* (pp. 1-42). Elsevier, Inc.: Academic Press.

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.*

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. Madison, WI: Cognitive Science Society.

Kominsky, J. F. and Phillips, J. (2019). Immoral Professors and Malfunctioning Tools: Counterfactual Relevance Accounts Explain the Effect of Norm Violations on Causal Selection. *Cognitive Science, 43*: e12792. doi:10.1111/cogs.12792

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

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. *Moral psychology*, *2*, 441-8.

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

Livengood, J., Sytsma, J., and 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.

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.

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.

Schaffer, J. (2013). Causal contextualisms. In Martijn Blaauw (Ed.) *Contrastivism in Philosophy*. Routledge, 43-71.

Sytsma, J. (2019). *Structure and Norms: Investigating the Pattern of Effects for Causal Attributions.* [Preprint]

Sytsma, J., & Livengood, J. (2019). Causal Attributions and the Trolley Problem. Preprint.

Sytsma, J., Livengood, J., and 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.