Marbles in Inaction: Counterfactual Simulation and Causation by Omission

Simon Stephan[1] (sstepha1@gwdg.de) Pascale Willemsen[2] (pascale.willemsen@rub.de)
Tobias Gerstenberg[3] (tger@mit.edu)
[1]Department of Psychology, Georg-August-University Göttingen and Leibniz ScienceCampus Primate Cognition
[2]Institute for Philosophy II, Ruhr-University Bochum
[3]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

Abstract

Consider the following causal explanation: The ball went through the goal because the defender didn't block it. There are at least two problems with citing omissions as causal explanations. First, how do we choose the relevant candidate omission (e.g. why the defender and not the goalkeeper). Second, how do we determine what would have happened in the relevant counterfactual situation (i.e. maybe the shot would still have gone through the goal even if it had been blocked). In this paper, we extend the counterfactual simulation model (CSM) of causal judgment (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014) to handle the second problem. In two experiments, we show how people's causal model of the situation affects their causal judgments via influencing what counterfactuals they consider. Omissions are considered causes to the extent that the outcome in the relevant counterfactual situation would have been different from what it actually was.

Keywords: causality; counterfactuals; causation by omission; causal attribution; mental simulation.

Introduction

Billy is on his way home. He is driving on a lonely country road, when he notices a damaged car next to the road. The car seems to have collided with a tree, and the driver appears unconscious. Billy decides not to stop and keeps driving. A few days later, Billy reads in the newspaper that the driver died because he had not received any medical attention.

Many people would concur that Billy's not having stopped was causally relevant for the driver's death. However, there are two fundamental problems with citing omissions (i.e., events that did not happen) as causes. First, there is the *problem of causal selection*. Why cite Billy's not stopping as causally relevant for the driver's death? Why not cite the Queen of England? Second, there is the *problem of underspecification*. Assuming that Billy would have stopped to check on the driver, what would he have done? Would Billy's acting have prevented the driver's death, or would she have died anyway?

In this paper, we show how the *counterfactual simulation model* (CSM) of causal judgment developed in Gerstenberg, Goodman, Lagnado, and Tenenbaum (2012) (see also Gerstenberg et al., 2014; Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2015) provides a natural solution to the underspecification problem. The CSM predicts that an omission is a cause when the positive event that is chosen as its replacement would have changed the outcome of interest. More specifically, we show how people's causal model of a situation guides their selection of the relevant counterfactual which subsequently determines their judgment about whether the omission made a difference to the outcome.

The paper is organized as follows: We first describe the causal selection and the underspecification problem in more detail. We then propose an extension to the CSM as a solution to the underspecification problem. Thereafter, we present and discuss the results of two experiments which test the CSM.

The Causal Selection Problem

Many philosophers argue that counterfactual approaches to causation are too inclusive when it comes to omissions (e.g. McGrath, 2005). If Billy had stopped and checked on the unconscious driver, the driver would not have died. Consequently, the driver died because Billy did not stop. However, following this logic, the same counterfactual seems to be true for the Queen of England. If the Queen of England had stopped, the driver would not have died either. However, intuitively, it is Billy's omission that was causally relevant, and not the Queen's. The problem of causal selection has been intensively discussed in both philosophy and empirical studies (e.g. Hesslow, 1988). Interestingly, while the causal selection problem presents a challenge to certain philosophical theories of causation, laypeople do not have any difficulty in selecting the cause of the driver's death. Based on evidence from research on causal cognition, it has been suggested that the concept of causation is not a purely descriptive one, but that it depends on reasoners' expectations (Willemsen, 2016). While we would have expected Billy to stop and help, we didn't entertain any such expectation for the Queen.

The Underspecification Problem

When it comes to omissive causation a fundamental problem is how to define the relevant counterfactual contrast (cf. Schaffer, 2005). For positive events ("something happened"), the counterfactual contrast ("it didn't happen") is often well-defined. However, replacing a negative event with a positive event seems more problematic because there are a infinitely many ways in which events can come about. If Billy actually helped the driver, it seems to be pretty clear what would have happened if he had not helped (he would just have continued to drive on). However, if Billy did not help, it is unclear what would have happened if he had helped (would he have helped in a competent manner to prevent the driver's death, or would he have been too nervous and screwed things up?).

While the causal selection problem has received much attention in the literature (e.g., Henne, Pinillos, & De Brigard, 2015; Livengood & Machery, 2007), the underspecification problem has not. One exception is the account by Wolff, Barbey, and Hausknecht (2010) that addresses both problems. The general idea proposed by Wolff et al. (2010) is that cau-

sation by omission is linked to the removal of an actual (or anticipated) force that previously prevented a certain outcome from occurring. One problem of this account, however, is that it appears too restrictive in that it cannot account for cases in which no (apparent) force is removed. Imagine, for instance, sentences like "The lack of rain caused the drought in Somalia". Here, it would be a stretch to think of a the lack of rain as the removal of a force.

The extension of the CSM that we propose in this paper provides a different solution to the *underspecification problem*. Previous research has suggested that the extent to which a certain counterfactual is relevant is a function of both how likely we are to consider it, and how likely it would have changed the outcome of interest (Petrocelli, Percy, Sherman, & Tormala, 2011). However, while this research has shown that these counterfactual probabilities affect people's causal judgments, it doesn't explain how we come up with the relevant probabilities in the first place. Here, we will show how the CSM provides a natural solution to determine whether an omission made a difference to the outcome.

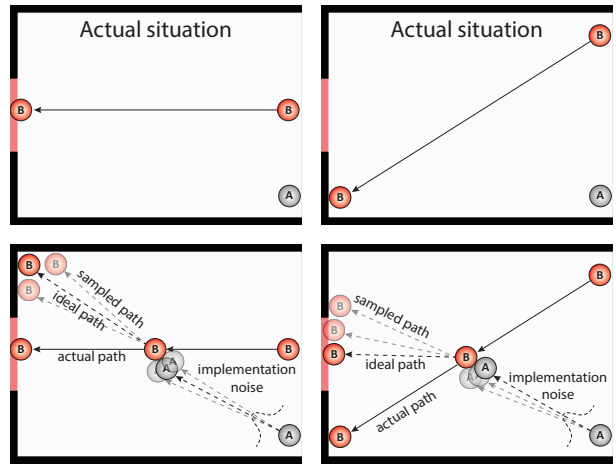
Counterfactual Simulation and Omission

The CSM predicts that people make causal judgments by comparing what actually happened with the outcome of a counterfactual simulation. So far, the model has been applied to capturing participants' judgments about events that actually happened (Gerstenberg et al., 2012, 2014, 2015). Consider the situation shown in Figure 1b (bottom) illustrated as the *ideal path*. Here, A collides with B and B subsequently goes through the gate. The CSM says that ball A's colliding with ball B caused ball B to go through the gate in this case, because it is obvious that ball B would have missed the gate but for the collision with A. More generally, the CSM predicts that causal judgments are a function of the reasoner's *subjective degree of belief* that the candidate cause made a difference to the outcome. More formally, we can express the degree of belief that x caused y as

$$P(x \triangleright y) = P(y' \neq y | \mathcal{S}, \mathrm{do}(x')), \tag{1}$$

in which x denotes the event of ball A hitting ball B, and the outcome y captures the event of ball B going through the gate. We first condition on what actually happened $\mathcal{S}$ (i.e., the motion paths of each ball, the position of the walls, etc.). We then intervene to set the candidate cause event x to be different from what it was in the actual situation, $\mathrm{do}(x')$. Finally, we evaluate the probability that the outcome in this counterfactual situation y' would have been different from the outcome y that actually happened. The results of several experiments (cf. Gerstenberg et al., 2012, 2014, 2015) have revealed that there exists a tight relationship between the counterfactual judgments of one group of participants (about what would have happened if the candidate cause had been absent), and the causal judgments of another group of participants.

To model causal judgments about positive events, the CSM considers counterfactuals in which the positive event (ball A colliding with B) is simply removed from the scene (indicated



(a) Did B **go through** the gate because A did not hit B?

(b) Did B **miss** the gate because A did not hit B?

Figure 1: Illustration of what actually happened (top) and the counterfactual simulation model (bottom). The diagrams illustrate the actual path that ball B took, as well as an ideal path for (a) A preventing B from going through the gate, or (b) A causing B to go through the gate. The sampled paths show example simulations that result from applying implementation noise to the ideal path. *Note*: In (a), A would have prevented B from going through the gate for both sampled paths. In (b), A would have caused B go through the gate in one sample but not so in the other in which B would still have missed even though A hit B.

by $\mathrm{do}(x')$ in Equation1). Things become more intricate, however, when we want to model omissions as causes. As discussed above, it is often straightforward to replace an event with a non-event (e.g., a collision with no collision), but it is less clear how to replace a non-event with an event. Consider the situation shown in Figure 1a. Did ball B go through the gate because ball A *did not* hit it? The problem is that there are infinitely many ways for ball A to collide with ball B. Which of these events are we to consider? The collision event is severely underspecified. We will now show how the CSM can be extended to yield predictions about omissions as causes, and thereby provide a solution to the underspecification problem.

Modeling Omissions

We assume that people solve the underspecification problem by sampling counterfactual possibilities based on their intuitive understanding of the situation (cf. Kahneman & Tversky, 1982). The extent to which the omission is viewed as a cause of the outcome is assumed to be a function of the proportion of samples in which the outcome would have been different from what actually happened, assuming that the type of counterfactual event of interest was realized. Let us illustrate how the model works by example of the situation depicted in Figure 1a. In the actual situation, ball A did not move and ball B went right through the middle of the gate. We want to determine to what extent A's not hitting ball B was a cause of B's going through the gate. To do so, we simulate what would have happened if ball A had collided with B. More specifically, we need to determine the time t at which A would have started to move, the direction d in which ball

A would have moved, and the velocity v. Once we have determined these quantities, we can simulate what would have happened. For many combinations of values for t, d, and v ball A would not have collided with ball B. We can discard all such situations since we are interested in evaluating what would have happened if ball A had *hit* ball B. For each situation in which the two balls collide, we record what the outcome would have been – would B have missed the gate, or would it still have gone through the gate? We can now obtain the probability that ball A's not hitting ball B was a cause of ball B's going through the gate (cf. Equation 1) by looking at the proportion of samples in which B would have missed the gate instead of going through.

But how do we determine what values to take for t, d, and v which jointly determine what counterfactual situation we consider? We predict that prior expectations guide the counterfactuals we consider. In Experiment 1 below, we contrast situations in which participants don't have any expectations about what normally happens, with situations in which participants have statistical, or social expectations. We will now discuss how the model incorporates these expectations.

Expectations Shape Counterfactual Simulations

No Expectations Let us first assume a situation in which an observer does not have any strong expectations concerning how the balls typically move in the given context. When asked whether A's not hitting B caused B to go through the gate, we have to generate situations in which A would have hit B. This already considerably constrains what kinds of situations we consider. For example, it would be futile to consider situations in which A only starts moving after B already went through the gate, or in which A moved toward the right.

We generated counterfactual samples in the following way: We first discretized the space for the time at which A starts moving t, the direction in which it moves d, and its velocity v. For t, we considered all values from 0 to $t_{outcome}$ where 0 corresponds to the time at which B starts moving and $t_{outcome}$ to the time at which ball B went through the gate (or hit the wall). For d, we considered the full range from A going straight to the left to going straight up. For v, we considered a reasonable range from A moving slowly to A moving fast. For each generated world, we noted whether A and B collided, and whether B went through the gate or missed the gate. We then discarded all situations in which the two balls did not collide, and recorded the proportion of situations in which B would have gone through the gate if the balls had collided.

The model makes the following predictions: For the situation in which B is on a path toward the gate (Figure1a), there is a good chance that B would have missed the gate if ball A had hit it. The model predictions are shown in Figure2. As can be seen in the left panel, the CSM concludes that the probability that B would have missed the gate had A hit it is just as high as the probability that B would have passed the gate. By contrast, when B is on a path away from the gate ("missed" in Figure 2, cf. Figure 1b top right) there is only a relatively small chance that ball B would have gone through the gate if ball A had hit it. Thus, the CSM predicts that people will be more likely to agree that ball B *went through* the gate because ball A did not hit it than they will be to agree that ball B *missed* the gate because ball A did not hit it.

Social Expectations When nothing particular is known about how A and B typically move, the space of counterfactuals from which the CSM samples is relatively wide. It seems plausible, however, that what counterfactual possibilities are considered will be affected by different forms of prior expectations. Imagine, for example, that you learn that two players play a marble game. Player B wants to get her marble into the goal, while Player A wants to make sure that this does not happen. On a particular trial, Player A did not pay attention and forgot to flick his marble. Did Player B's marble go through the gate because Player A's marble did not hit it? When knowing that it is a player's job to prevent a marble from going through the gate, people may expect that this player would not have just flicked her marble randomly. Instead, she can be expected to try her best to make sure that the other marble does not go through the gate. Similarly, consider a situation in which Player A also wants that Player B's marble goes through the gate. In that case, it seems likely that Player A will try to flick his marble so that it makes sure that B's marble will go through the gate.

Figure 1 illustrates how the CSM incorporates how prior expectations constrain the space of counterfactual situations. We assume that the player would first determine a time t at which to flick her marble. For any given point t, the player then determines an optimal d and v conditional on the player's goals. For a player who wants to prevent ball B from going through the gate, the player's goal is to maximize the distance between B's position and the middle of the gate. For a player who wants to cause B to go through the gate, the player's goal is to minimize the distance between B's position and the middle of the gate (i.e., she wants B to go right through the middle of the gate). For simplicity, we assume that players can plan their action optimally, but that they have some implementation noise. The CSM models this implementation noise by introducing a small perturbation to the ideal path on which A moves. As is illustrated in Figure1, the CSM incorporates implementation noise by slightly perturbing the "ideal path" vector.

Figure 1 shows the actual path that ball B took, the ideal paths that player A "wanted" the marbles to take, and two examples for paths that ball B actually took after subjecting A's ideal plan to some implementation noise. Notice that the implementation noise has a larger effect in Figure 1b where it leads to a situation in which ball B would have missed the gate even though ball A hit it. In contrast, in Figure 1a the implementation noise has less of an effect. Here, ball B would reliably miss the gate even if we apply some implementation noise to player A's intended plan. Accordingly, the CSM predicts that it is more likely that A's hitting B would have resulted in B missing the gate (when B actually went through,

Figure 1a) than it would have resulted in B going through the gate (when B actually missed, Figure 1b). Since the sample of considered situations is biased toward optimal actions, the CSM predicts that judgments will overall be higher than when an observer does not have any prior expectations. The predictions for this situation are shown in the middle panel in Figure 2.

Statistical Expectations Now imagine that instead of learning anything about agents playing a game you get to see a few situations first that shape your expectations about what tends to happen. We incorporate such "statistical" expectations into the model in the same way in which we handled social expectations. However, we allow for the implementation noise to be different between these situations. Specifically, the size of the implementation noise parameter will depend on the kind of evidence that participants have seen. For example, if one has witnessed a series of trials in which A always hit B in such a way that B went straight through the gate, this would suggest a smaller implementation noise compared to one that is suggested by trials in which A hit B in such a way that B went through the gate in, for example, merely two third of the cases. The predictions for this situation are shown in the right panel in Figure 2.

Experiment 1

Experiment 1 tests whether the CSM accurately predicts people's causal judgments for omissions in dynamic physical scenes. We look at causal judgments about situations in which ball A failed to hit ball B, and ball B either went through or missed the gate (see Figure 1). In line with the CSM, we predict that the degree to which people judge ball A's not hitting ball B as causally relevant to the outcome would be tightly coupled with the results of a mental simulation about what would have happened if a collision had occurred. Furthermore, we test the hypothesis that different types of expectations (social or statistical) influence people's causal judgments by affecting what counterfactual situations people consider.

Methods

Participants and Materials 476 participants (239 female, $M_{Age} = 33.83$ years, $SD_{Age} = 12.03$ years) were recruited via Prolific Academic (`www.prolific.ac`) and participated in this experiment for a monetary compensation of £0.25. The clips were created in Adobe Flash CS5 using the physics engine Box2D.

Design and Procedure All factors were manipulated between subjects. We manipulated what actually happened (*actual outcome*: missed vs. went through), and the expectations of participants about what will happen (*expectation*: no expectations, statistical expectation, social expectation). Finally, we varied whether participants answered a causal question, or a (counterfactual) probability question (*question*: causation vs. probability).

In the "no expectations" condition, subjects simply read that they will see an animation in which a stage with solid walls, two balls A and B, and a gate will be displayed. All subjects were shown a graphical illustration of the stimuli. Participants in the "statistical expectation" condition were presented four primer clips in which ball B actually collided with A. One group of subjects saw that the collision always *caused* B to go through the gate, while the other half always saw that A *prevented* B from going through the gate (see Figure 1). In the "social expectation" condition, subjects were instructed that the video clip (which was the same as in the "no expectations") shows what happened during a game of marbles played by two agents, Andy and Ben. We manipulated whether subjects believed that Andy wants to *help* Ben to flip his marble through the gate or whether he wants to *hinder* Ben from doing so.

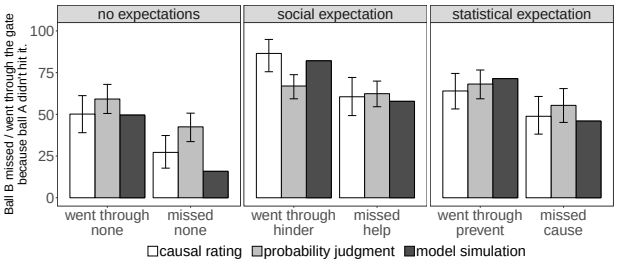


Figure 2: **Experiment 1**. Mean causal and probability judgments together with the predictions of the CSM. *Note*: Error bars indicate 95% bootstrapped *CI*s.

Participants in the "causation" condition indicated how much they agreed with the claim that B *missed* the gate because A did not hit it, or that B *went through* the gate because A did not hit it, depending on the outcome. Participants in the "probability" condition gave a corresponding probability judgment: they indicated what they believed the chances were that B would have gone through / missed the gate if ball A had hit ball B. Participants indicated their ratings on a sliding scale.

Which outcome participants saw depended on the expectation condition: In the "social expectation" condition, participants who expected the agent to *help* saw that B actually missed the gate, and participants who expected the agent to *hinder* saw that B went through the gate. In the "statistical expectations" condition, participants who had seen the *causation* clips saw that B missed the gate, whereas those who had seen the *prevention* clips saw that B went through the gate.

Results and Discussion

Figure 2 shows participants' mean causal ratings (white bars), probability ratings (gray bars), as well as the predictions of the CSM (black bars). The CSM correctly predicts a difference in agreement ratings for both the causal and probability condition as a function of the outcome (went through vs. missed). A global 2 (question) $\times$ 6 (combination of expectation and outcome) factorial ANOVA shows a main effect of outcome, $F(5,464) = 14.51$, $p < .001$, $\eta^2_G = .61$ but no main effect of question, $F(1,464) < 1$. The interaction between question and expectation was significant,

$F(5,464) = 2.74$, $p < .05$ but the effect is small, $\eta^2_G = .03$.

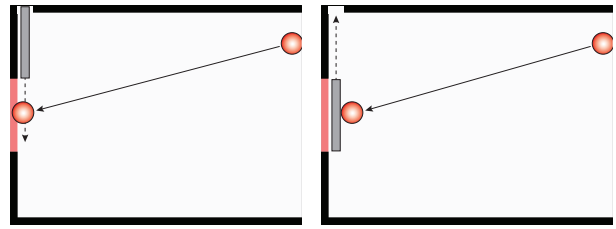
Importantly, participants saw A's not hitting ball B as more causal when B went through the gate compared to when it missed. This pattern was predicted by the CSM and indicates that participants' counterfactual simulations and their causal inferences were sensitive to the constraints imposed by the virtual physical environment. Because the displayed gate was relatively small, the probability that a collision would change the outcome is higher if B actually went though, than when it missed. Planned contrasts confirmed that the observed differences between "went through" and "missed" were significant in all conditions, with $t(464) = 3.21$, $p < .01$, $r = .15$ in the "no expectations" condition, $t(464) = 2.13$, $p < .05$, $r = .10$ in the "statistical expectation" condition, and $t(464) = 3.53$, $p < .001$, $r = .16$ in the "social expectation" condition.

Besides the asymmetry between "went through" and "missed", we also expected to see higher causality ratings in the "statistical" and "social expectation" conditions than in the "no expectation" condition. This difference was predicted because we incorporated an ideal path in these situation that was then perturbed by imposing some implementation noise. As Figure 2 shows, we did indeed observe this pattern. A planned contrast confirmed that this difference was significant, $t(464) = 5.98$, $p < .001$, $r = .27$.

Concerning the probability ratings, planned contrasts showed that the difference between "went through" and "missed" was significant in the "no expectations condition", $t(464) = 2.33$, $p < .05$, $r = .11$, and the "statistical expectation" condition, $t(464) = 1.73$, $p < .05$, $r = .08$, but not in the "social expectation" condition, $t(464) < 1$. Concerning the predicted difference between the "no expectations" condition and the other two expectation conditions, Figure 2 shows that we obtained a similar pattern as for the causality judgments. In line with our expectations, the probability ratings for the "statistical expectation" and the "social expectation" condition were higher than the ratings for the "no expectation" condition, $t(464) = 2.82$, $p < .01$, though this effect was smaller than the effect for the causality judgments, $r = .13$.

The results of Experiment 1 show that participants' causal judgments are qualitatively well accounted for by the CSM. The CSM also does a good job in accounting for the pattern quantitatively, as evidenced by a high correlation between model predictions and counterfactual probability judgments ($r = .97, RMSE = 14.00$), as well as between model predictions and causal judgments ($r = .97, RMSE = 6.06$). The fact that the model accounts slightly less well for the counterfactual probability judgments is mainly due to the relatively large difference between model predictions and probability judgments in the "no expectations" condition.

A key finding in Experiment 1 is the asymmetry in participants' causal judgments as a function of whether ball B went through or missed the gate. The CSM predicts this pattern because it is more likely that A's hitting B would prevent B from going through the gate (cf. Figure 1a) than that it would cause B to go through (cf. Figure 1b). One possibility, however, that Experiment 1 cannot rule out is that people are in general more likely to regard omissions as causes when the relevant counterfactual involves preventing compared to causing. In Experiment 2, we investigate whether there is such a general asymmetry between omissive causation and prevention.



(a) Did the ball **go through** the gate because the wall did not move? ($M = 87.51$, $\pm 95\%$ $CI = 7.67$)

(b) Did the ball **miss** the gate because the wall did not move? ($M = 89.00$, $\pm 95\%$ $CI = 8.37$)

Figure 3: Illustration of the materials used in Experiment 2. Solid arrows indicate the actual path of the ball; dashed arrows show the hypothetical path of the wall. Graph (a) shows the "went through" and (b) the "missed" condition. The results for both conditions are included in brackets.

Experiment 2

The goal of Experiment 2 was to rule out that the observed difference between "went through" and "missed" in Experiment 1 came about because people generally treat omissive causation and omissive prevention differently. The CSM only predicts an asymmetry between two situations when the positive event of interest was more likely to make a difference in one situation compared to the other. Hence, our strategy in Experiment 2 was to hold this probability constant. To achieve this goal, we simply replaced ball A with a wall that had exactly the size of the gate. To model "missed" and "went through", we varied whether the wall blocked the gate or not, while a displayed ball always headed toward the gate (see Figure 3). Participants rated how much they agree that "the ball" missed the gate (or went through the gate) because the wall did not move. There is no ambiguity about the relevant counterfactual in this case – it is clear that the outcome would have been different, had the wall moved. Accordingly, the CSM predicts that participants' judgments should be high for both cases, no matter whether the ball went through the gate or missed the gate because of the omission.

Method

Participants 65 participants (40 female, $M_{age} = 32.86$, $SD_{age} = 12.84$) who were again recruited via Prolific Academic completed this online experiment and received a monetary compensation of £ 0.25.

Design, Materials, and Procedure The final outcome, that is, whether the ball went through or missed the gate (see Figure 3) was manipulated between subjects. The instructions were similar those used in the "no expectations" condition in Experiment 1. Further, participants were presented an illustration showing the materials in which it was made clear that the wall can only be in two different positions, either right

in front of the gate or in the upper left corner of the stage (see Figure 3). Having read the instructions, participants were shown the respective video clip and provided the causal rating after the clip was finished.

Results and Discussion

As expected, participants gave very high causal ratings for "went through" ($M = 87.51$, $SD = 21.62$) and "missed" ($M = 89.00$, $SD = 23.21$). As predicted by the CSM, the ratings were not different from each other, $t(63) < 1$. The probability that the outcome would have been different in the relevant counterfactual, is close to maximal in both conditions.

The results of Experiment 2 are in line with the CSM. Further, the fact that the causality ratings were both very high and not different from each other rules out the potential alternative explanation that people might generally treat omissive causation and omissive prevention differently.

General Discussion

We developed an extension of the *Counterfactual Simulation Model* to account for causation by omission. Based on previous research by Gerstenberg et al. (2014), we reasoned that people's causal judgments are closely linked to their subjective degree of belief that the outcome would have been different had the candidate cause been replaced. We argued that this replacement by a counterfactual contrast is particularly difficult in cases of omissions. The counterfactual contrast to "did not hit" is clearly "had hit", but it remains unclear what would have happened if "hitting" had taken place.

In two experiments we shed light on how to tackle the underspecification problem. We predicted that prior expectations would constrain what counterfactual contrasts people consider relevant to the scenario. Experiment 1 revealed an asymmetry: A's not hitting B was judged less causal when B missed the gate compared to when B went through the gate. This is what the CSM predicts, and the results thus lend additional support to the hypothesis that causal judgments are grounded in counterfactually simulated probabilities. Adding expectations increased both people's causal judgments as well as their subjective degree of belief that a counterfactual collision would changed the outcome. This effect was particularly strong for social expectations, which the CSM explains by assuming that knowledge about intentions of agents limits the range of counterfactuals that are considered. Our results thus add to previous research indicating that intentional actions signal higher causal stability compared to unintentional ones (Lombrozo, 2010), and that causal stability is indeed a relevant dimension that affects causal reasoning (Nagel & Stephan, 2016).

It might be objected that the asymmetry in causal attribution for "went through" and "missed" in Experiment 1 is not due to a difference in what would have happened in the relevant counterfactual simulations, but rather due to an inherent asymmetry between omissions that prevent and omissions that cause. Experiment 2 addressed this possible confound by looking at situations in which the relevant counterfactual event was clear (a wall that could only move in one direction), as well as what would have happened in case that event had happened. Just as predicted, we found that causal ratings were equally high irrespective of whether the ball "went through" and "missed" in this case. Instead of a general asymmetry between prevention and causation, participants judge omissions to be causal the more certain they are that the omission made a difference to the outcome.

As our introductory example demonstrates, omissions are particularly relevant in human interaction, especially so in morally or legally charged situations when we had clear expectations about what a person should have done. In this paper, we have shown how the CSM accounts for people's causal judgments of omissions in situations in a physical domain in which the relevant counterfactuals are relatively well constrained. However, we believe that the CSM has the potential to capture causal judgments about omissions of social agents as well. For example, the extent to which we blame someone for not having helped depends on how easy it would have been for the agent to help (cf. Jara-Ettinger, Tenenbaum, & Schulz, 2015). In future research, we will explore the CSM in a richer social setup.

Acknowledgments This project was supported by the Leibniz Association through funding for the Leibniz ScienceCampus Primate Cognition. TG was supported by the Center for Brains, Minds & Machines (CBMM), funded by NSF STC award CCF-1231216.

References

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2012). Noisy Newtons: Unifying process and dependency accounts of causal attribution. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 378–383). Austin, TX: Cognitive Science Society.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2014). From counterfactual simulation to causal judgment. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 523–528). Austin, TX: Cognitive Science Society.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2015). How, whether, why: Causal judgments as counterfactual contrasts. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 782–787). Austin, TX: Cognitive Science Society.

Henne, P., Pinillos, Á., & De Brigard, F. (2015). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, 1–14.

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). Brighton, UK: Harvester Press.

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not so innocent: Toddlers' inferences about costs and culpability. *Psychological Science*, *26*(5), 633–640. Retrieved from `http://dx.doi.org/10.1177/0956797615572806` doi: 10.1177/0956797615572806

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Livengood, J., & Machery, E. (2007). The folk probably don't think what you think they think: Experiments on causation by absence. *Midwest Studies in Philosophy*, *31*(1), 107–127.

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

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

Nagel, J., & Stephan, S. (2016). Explanations in causal chains: Selecting distal causes requires exportable mechanisms. In A. Papafragou, D. Grodner, D. Miram, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 806–812). Austin, TX: Cognitive Science Society.

Petrocelli, J. V., Percy, E. J., Sherman, S. J., & Tormala, Z. L. (2011). Counterfactual potency. *Journal of Personality and Social Psychology*, *100*(1), 30–46.

Schaffer, J. (2005). Contrastive causation. *The Philosophical Review*, *114*(3), 327–358.

Willemsen, P. (2016). Omissions and expectations: A new approach to the things we failed to do. *Synthese*. Advance online publication. doi: 10.1007/s11229-016-1284-9

Wolff, P., Barbey, A. K., & Hausknecht, M. (2010). For want of a nail: How absences cause events. *Journal of Experimental Psychology: General*, *139*(2), 191–221.