**The tossing sea of Cause and Theory: Using eye movements to adjudicate theories of causal selection**

Kristina Krasich1, Kevin O’Neil1,2, Felipe De Brigard1,2,3,4

[kristina.krasich@duke.edu](mailto:kristina.krasich@duke.edu), [kevin.oneill@duke.edu](mailto:kevin.oneill@duke.edu), [felipe.debrigard@duke.edu](mailto:felipe.debrigard@duke.edu)

|  |  |
| --- | --- |
| 1 Center for Cognitive Neuroscience  Duke University  Durham, NC, USA | 3 Duke Institute for Brain Science  Duke University  Durham, NC, USA |
| 2 Department of Psychology and Neuroscience  Duke University  Durham, NC, USA | 4 Department of Philosophy  Duke University  Durham, NC, USA |

Corresponding author:

Kristina Krasich

Center for Cognitive Neuroscience

Duke University

Durham, NC, USA

[kristina.krasich@duke.edu](mailto:kristina.krasich@duke.edu)

**Word Count: XX,XXX**

Author note: [contributions, thanks, and…] This project is supported by an Office of Naval Research award (N00014-17-1-2603) to FDB.

**Abstract**

*Keywords:* Causal selection, counterfactual thinking, mental simulation, eye movements

**Introduction**

Whether with intentional deliberation or spontaneous intuition, people are frequently engaged in causal reasoning. Causal selection—which includes evaluating causal candidates and determining which, among them all, was the actual cause—in particular allows people to modify their future behaviors, avoid harmful outcomes, and achieve optimal ones. How people select causal agents, however, remains unclear, given that such reasoning is highly influenced by a variety of factors, such as perceived norms and expectations (e.g., Knobe, 2009; Knobe & Fraser, 2008), the degree to which candidate causes and effects covary (Cheng & Novick, 1991; Wasserman, Chatlosh, & Neunaber, 1983), events that did not happen (e.g., Henne, Niemi, Pinillos, De Brigard, & Knobe, 2019; Clarke, Shepherd, Stigall, Waller & Zarpentine, 2015), and causal structure (e.g., Icard, Kominsky, & Knobe, 2017). Inaccurately selecting causal agents presents potentially severe legal and/or social ramifications (Hart & Honoré, 1985), so to avoid such consequences, it is critical to understand the process by which people navigate the complexities of causal selection.

**Mechanistic accounts of causal reasoning**

*Process theories* of causal selection argue that people will focus on the direct interaction between covarying events to assess whether there was an exchange of some conserved quantity, such as energy or momentum, from the candidate cause to the given outcome (Dowe, 2000; Salmon, 1997; Wolff, 2007). Supportive evidence shows a greater propensity for people to attribute causation when there is a continuous mechanism that links an action (e.g., kicking a ball) to the outcome (e.g., the ball landing in the goal; Walsh & Sloman, 2011). Relatedly, events that are identified as candidate causes usually directly precede the outcome, and possible enabling conditions (e.g., effective coaching strategies) are less often judged as candidate causes (e.g., Goldvarg & Johnson-Laird, 2001). Therefore, process theories argue that a single consideration of what *actually happened* between the candidate cause and the outcome will be sufficient, and preferred, for establishing causal selection.

As one alternative account, *counterfactual theories* argue that people engage in counterfactual thinking to select casual agents (e.g., Lewis, 1973). That is, they consider alternative scenarios to assess whether the candidate cause made a difference in the actual outcome. Counterfactual thinking can be highly episodic, with memory-like spatiotemporal episodes that resemble, but are altered from, the actual autobiographical event (Roese & Epstude, 2017). Counterfactual thinking, thus, engages cognitive (e.g., Weisberg & Gopnik, 2016) and neural (e.g., Parikh, Ruzic, Stewart, Spreng, & De Brigard, 2018; Kulakova, Aichhorn, Schurz, Kronbichler, & Perner, 2013) processes that support both memory reactivation as well as hypothetical simulation (for review, see De Brigard & Parikh, 2019). More nuanced accountsshow that counterfactual thinking frequently occurs when outcomes are negative and/or unexpected (Roese, 1997), but that the counterfactual alternative needs to be relevant (Halpern & Hitchcock, 2014; Icard et al., 2017; Kominsky & Phillips, 2018) and plausible (Petrocelli, Percy, Sherman, & Tormala, 2011) to impact causal judgments. Therefore, the idea underlying counterfactual theories is that people will consider what actually happened along with to what reasonably *might have happened* had the candidate cause been absent or altered in some way.

Gerstenberg et al., (2017) sought to adjudicate between process andcounterfactual theories by using eye movements to help characterize the cognitive operations engaged during casual selection. This approach was empirically grounded in a large body of research showing a tight relationship between gaze behaviors and the real-time visuospatial information-processing priorities of the visual system (e.g., Just & Carpenter, 1976; Kowler et al., 1995). That is, eye movements reveal what visuospatial information is overtly attended, encoded, and/or recalled, at a given moment in time. Gerstenberg et al., (2017) measured eye movements of participants watching videos of two balls (Balls A and B) moving toward a goal and colliding into each other. After the collision, Ball B’s heading changed such that it sometimes missed the goal, and it sometimes scored. As a between-subject condition, participants were instructed to engage in one of three possible thoughts while watching the video: 1) assess the extent to which Ball B did or did not score into the goal (outcome assessment), 2) think about what would have happened to Ball B had Ball A not been present (counterfactual thinking) or, 3) judge whether Ball A colliding into Ball B caused or prevented Ball B from scoring in the goal (causal selection). The authors then compared gaze behaviors across these three conditions to gauge whether the gaze behaviors of participants who engaged in causal selection were more similar those who engaged in counterfactual thinking or outcome assessment.

The findings showed that the participants who engaged in causal selection and those who engaged in counterfactual thinking showed similar gaze behaviors. That is, these participants more frequently fixated where Ball B would have traveled had Ball A not interfered. This tendency also corresponded with Ball A being judged as more causally related to the outcome of Ball B. In comparison, participants who engaged in outcome assessment tended to just look directly at Ball B. The authors therefore inferred that the participants who engaged in causal selection, to some degree, mentally simulated the counterfactual movements of Ball B to gauge the causal influence of Ball A.

There were several experimental limitations, however, that challenged this claim. First, the counterfactual movement of Ball B was spatially equivalent to its original projected heading prior to Ball A’s interference. This display layout made it impossible to disambiguate whether fixations to this location actually reflected counterfactual simulations or simply anticipatory, future-directed hypothetical simulations of Ball B’s original projected heading. The authors acknowledged this limitation but suggested that even if the eye movements did not actually reflect counterfactual thinking, fixations toward Ball B’s projected heading would still allow participants to acquire the same visuospatial information needed for generating counterfactual contrasts *after* the visual input was removed. This offered explanation suggests that how visuospatial information is initially encoded will impact how it is retrospectively considered for causal selection. This idea, though, was not tested.

As a further limitation, gaze behaviors in general are highly observer-driven, swayed by the observer’s prior knowledge and expectations (refs), short-term memory of previously fixated visuospatial information (refs), and the strategies evoked for achieving task-specific goals (refs). Measuring outcome assessment, counterfactual thinking, and causal selection as between-subject conditions was accordingly problematic given that participants could strategically—and uniquely—select in advance what visuospatial information to acquire and ignore. Without controlling for these factors within subjects, the unique impact of these different thought types on gaze behaviors was indeterminable. Considering these critical limitations, it remained unclear whether participants were indeed, to some degree, engaging in counterfactual thinking for causal selection.

**The current study: Approaches and hypotheses**

The current research aims to adjudicate between process and counterfactual theories of causal selection by investigating how visuospatial information is initially encoded as well as mnemonically reconstructed during retrospective outcome assessment, counterfactual thinking, and causal selection. In a mixed experimental design, participants will complete a ball-shooting-paradigm where they will try to shoot a ball into a goal that is guarded by a goalie. Participants will then retrospectively reflect on the outcome and answer percept-related questions, including questions regarding the causal relevance of the ball or the goalie on the outcome. Each of these steps are illustrated in **Figure 1A** and described next.

*Attempting to score and encoding the outcome*

Participants will first decide whether to shoot a ball to the left or right of a goal in an attempt to score (**Figure 1A**, **Decision**). While their eye movements are recorded, participants will then watch a video of the outcome to encode whether the ball successfully scored or whether a computer-controlled goalie blocked it (**Figure 1A, Outcome video**). Specifically, as illustrated in **Figure 1B**, the ball will move diagonally in the chosen direction toward the goal. Meanwhile, a rectangular goalie will move horizontally either left or right. The goalie will block the ball from scoring if they move in the same direction, but the ball will score if they move in opposite directions. This display layout clearly differentiates the projected heading of each object from their counterfactual movements. For instance, if the ball moves to the right, its counterfactual movement is to the left. In this way, gaze behaviors oriented to counterfactual movements will be distinct from any anticipatory, future-oriented hypothetical simulations of the objects’ projected heading.



**Figure 1.** A) Example trial sequence. Participants will first decide whether to shoot the ball to the left or to the right of the goal (Decision prompt). They will then watch a video of the outcome (Outcome video) to encode whether the ball scored or was blocked by the goalie. The example illustrated here includes the upward orientation. After the outcome video, a word will then appear in the center of the screen (Thought prompt) indicating how participants should then retrospectively reflect on the outcome (Reflection). Last, participants will provide percept-related judgements about the outcome (Judgements). B) Example outcome video orientations. In the illustrated upward orientation, the ball and the goalie are projected to move to the right, and their counterfactual movements would be to the left. In the illustrated downward orientation, the ball and goalie are projected to move to the left, and their counterfactual movements would be to the right. In both examples, the goalie would block the ball from scoring because they moved in the same direction (miss trial).

As between-subject conditions, participants will be asked to encode the outcome focusing on either the ball (*ball-focused*) or the goalie (*goalie-focused*). Because participants will choose the ball’s but not the goalie’s heading, this manipulation creates personal and impersonal viewpoints. Personal and impersonal episodic simulations, including counterfactual ones (De Brigard, Spreng, Mitchell, & Schacter, 2015), engage similar but dissociable neural processes (Addis, Wong, & Schacter, 2007; Addis et al., 2009; Faul, Jacques, DeRosa, Parikh, & De Brigard, in press; Hassabis et al., 2007). Moreover, people tend to attribute successes to personal factors and failures to impersonal ones (e.g., Bernstein, Stephan, & Davis, 1979), and these responsibility attributions are related to causal selections (e.g., Phillips & Shaw, 2014). As suggested by this past research, manipulating an observer’s perspective will likely impact how the outcome is encoded, counterfactually simulated, and judged. For instance, participants focused on the ball might fixate on the ball, its projected heading, and—potentially—its counterfactual movement more frequently than the corresponding goalie-related visuospatial information. We predict that the tendency to do so will also correspond to higher judgments of the ball’s causal relevance. We further anticipate the opposite findings for those focused on the goalie. Moreover, manipulating participants’ perspective will ensure that any eye movements evoked during encoding will reflect deliberate, overt cognitive processes as opposed to any potential underlying inherent biases to look at visuospatial information related to one object over the other.

As within-subject conditions, participants will later reflect on the outcome either by engaging in outcome assessment, counterfactual thinking, or causal selection. Importantly, though, at initial encoding, participants will be unaware of which type of reflection they will later engage. This design thus requires participants to encode all of the potentially relevant visuospatial information to best prepare, and it controls for any idiosyncratic observer-related strategies for doing so. As such, we do not predict differences in eye movements at initial encoding across trials with different retrospective reflections. It is possible, though, that participants will look at the objects’ *counterfactual movements*—the spatial locations where the objects would have traveled had they moved in the opposite directions—in preparation for retrospective counterfactual thinking. This idea is consistent with the speculations from Gerstenberg et al. (2017) that looking at counterfactual movements allows for counterfactual contrasts after visual inputs are removed. Therefore, we will assess the extent to which participants look at counterfactual movements during initial encoding to disambiguate this possibility. If participants do indeed look at counterfactual movements during encoding, we further predict that this tendency will correspond with more extreme judgements of causal relevance of the ball or the goalie on the outcome. This finding would favor counterfactual theories of causal selection. If, however, a tendency to look at counterfactual movements at encoding does not predict causal relevance judgements, or if participants do not look at counterfactual movements during encoding, findings would align with process theories.

*Retrospectively reflecting on the outcome and gauging causal relevance*

After the video, a centrally presented word will appear briefly indicating how participants should reflect on the outcome (**Figure 1A**, Prompt). Then, looking at a blank screen (**Figure 1A**, Reflection), participants will retrospectively reflect on what just occurred (*outcome assessment*), a possible alternative outcome (*counterfactual thinking*), or the candidate cause of the outcome (*causal selection*). Following speculations from Gerstenberg et al., (2017), gaze behaviors during initial encoding allow for the acquisition of visuospatial information that is needed for making counterfactual contrasts for causal selection. Therefore, gaze behaviors will be recorded during retrospective reflections to reveal the overt mnemonic reconstruction of the previously encoded visuospatial information in the absence of ongoing visual inputs.

This approach is supported by research showing that during mental simulation people tend to systematically move their eyes to spatial locations once occupied by the original visual information (e.g., Altmann, 2004; Brandt & Stark, 1997; Richardson & Spivey, 2000). These gaze behaviors correspond with greater recall for the visual information, especially if they follow similar gaze patterns as those enacted during initial encoding (fixation reinstatement; e.g., Bone et al., 2019; Gurtner, Bishof, & Mast, 2019; Laeng & Teodorescu, 2002; Martarelli & Mast, 2013). Fixation reinstatement in particular is strongly correlated with the reported vividness of the mental images and to similarities in neural action patterns between initial encoding and later imagery (Bone et al., 2019). Indeed, mental imagery in general is known to recruit similar neural activation patterns as external visual perception (for reviews see Pearson et al., 2015; Pearson, 2019), and the reported vividness of such imagery is related to the extent of this neural overlap (Dijkstra, Bosch, & Gerven, 2017). These collective findings indicate that gaze behaviors, including fixation reinstatement, can delineate which components of a visual mental image are being simulated and reflect the vividness of such mental simulation.

In light of this past research, we will compare gaze behaviors—including fixation reinstatement—during retrospective outcome assessment, counterfactual thinking, and causal selection to differentiate what visuospatial information is mnemonically reconstructed and used for causal selection. If gaze behaviors for causal selection more closely reflect counterfactual thinking as opposed to outcome assessment, the evidence would support the idea that participants did indeed, to some degree, consider counterfactual movements for causal selection. We would then further predict that the tendency to look at counterfactual movements would correspond higher judgments of causal relevance. Conversely, if gaze behaviors of causal selection more closely resemble those observed during outcome assessment, and this tendency corresponds with higher judgments of causal relevance, the evidence would support process theories. The extent to which gaze behaviors during initial encoding are reinstated during retrospective reflection should correspond with the subjective vividness of the mental simulation (refs). Therefore, we will further explore how reported vividness might correspond with judgements of causal relevance (e.g., Swann & Miller, 1982).

*Judging the outcome*

As the final step for each trial, participants will answer a series of percept-related questions regarding the given outcome and retrospective mental simulation (**Figure 1A**, Judgements). Outcome assessment trials will inquire about the extent the ball scored; counterfactual thinking trials will inquire about the extent the ball would have scored had either the ball or the goalie moved in the opposite direction; and causal selection trials will inquire about the causal relevance of either the ball or the goalie. As previously described, we will compare the how gaze behaviors correspond with these judgements. Additionally, we will also use judgements from outcome assessment and counterfactual thinking trials to predict causal judgments. In particular, the extent to which judgements from counterfactual thinking predict causal judgements will suggest the extent to which counterfactual information is used for causal selection: a stronger predictive relationship will support counterfactual theories. Lastly, participants will also report on the subjective vividness of the mental simulation and, as previously described, these reports will be assessed with regard to the gaze behaviors evoked during retrospective reflections as well as causal judgements.

**Methods**

**Participants**

Because there is little consensus on the proper approach for conducting an a priori power analysis within a linear mixed-effects framework, we will first establish a target sample size of 40 participants, which was the sample size used in Gerstenberg et al. (2017). We will then over-recruit by ~10% to account for possible cancellations and technical issues, for a total estimate of 44 recruited participants. Further, we will conduct Monte Carlo simulations to estimate the minimal detectable effect size in our data, which is closely linked to the given power and sample size, and will interpret our results in the context of those analyses.

Participants will be volunteers recruited from Duke University and the local community.

They will be randomly assigned to either the *ball-focused* or the *goalie-focused* conditions*.* Informed consent will be obtained from each participant following procedures approved by the University Institutional Review Board, and participants will be compensated $12/hr.

**Stimuli and apparatus**

The stimuli will consist of video clips that will be generated with JBox2D. The videos will be presented centered on a screen with a refresh rate of 50 Hz. All stimuli will be presented on a 24-in LCD monitor with a screen refresh rate of 59 Hz. Viewing distances of 94-cm will be maintained with a desk-mounted chin and forehead rest. Therefore, the videos will subtend 13° x 10° of visual angle.

The videos will contain three objects that move around and interact. These objects (illustrated in **Figure 1B**) will include: 1) a red goal, in which the participants are trying to score, 2) a blue ball, which participants will decide where to shoot in attempt to score, and 3) a black, rectangular goalie, which will move horizontally left or right in attempt to block the ball. The ball will always start centered along the edge of the display, it will always move diagonally in the direction chosen by the participant, and it will always move at the same angle and speed. The goalie will always start in the middle of the goal, it will always move horizontally either to the left or the right at the same time as the ball, and it will always move at the same speed each trial. The orientation of the display will vary by 180° on half of trials, resulting in an upward and downward orientation, which will be randomized by block.

While participants view these videos, we will track eye movements using the EyeLink 1000 Plus (SR Research, Inc.), sampling at a rate of 1000 Hz. The eye-tracker will be calibrated using a nine-point calibration at the beginning of the study. A one-point calibration will be used before each video to correct for drift in eye tracking validity that may occur naturally over time. Saccades will be operationalized as changes in recorded fixation position that exceeds 0.2° with either a velocity that exceeds 30°/s or an acceleration that exceeds 9,500°/s2. All participant responses will be registered with a standard computer mouse click.

**Design and procedures**

After providing written consent and following the 9-point calibration procedures, participants will receive detailed instructions (**Appendix A**) as well as several instructional videos to learn how the objects can move and interact with each other. These videos will expose participants to the starting position of each object, the speed and angle by which the objects move, and how the ball may score or miss the goal according to whether the goalie blocked the ball.

Participants will then complete several tasks each trial, which are illustrated in **Figure 1A**. First, participants will decide whether to shoot the ball to the left or right of the goal in an attempt to score (**Figure 1A**, **Decision**). They will then watch a video of the outcome to encode whether the ball scored or missed. Unknown to the participants, *score* (50% of total trials) and *miss* (50% of total trials) trials will occur equally often, randomized. Participants will be asked to encode the video focused on the ball (*ball-focused* *condition*) or the goalie (*goalie-focused* *condition*).

Participants will then see (2 s) one following centrally presented prompts: Remember, What If, or Cause. If participants see the prompt *Remember*, they should think about the actual sequence of events that just occurred (*outcome assessment*). If participants see the phrase *What If*, they should think about what would have happened had the ball or the goalie moved in a different direction (*counterfactual thinking*). If participants see the word *Cause*, they should think about the candidate cause for the ball scoring or not scoring (*causal reasoning*). The instructions will

Last, at their own pace, participants will answer a series of reflection-specific questions about the given outcome, which are listed in **Table 1** and described below in sequential order. The next trial will start once all questions are completed. Participants will complete 4 blocks of 18 trials. All experiment procedures are estimated to take no more than 60 minutes.

*Table 1. Reflection-specific questions asked at the end of each trial*

|  |  |
| --- | --- |
| **Vividness of mental image** | |
| Outcome assessment | How vividly could you visualize what [your ball/the goalie] just did? |
| Counterfactual thinking | How vividly could you visualize what would have happened if [your  ball/the goalie] went in the other direction? |
| Causal selection | How vividly could you visualize whether [your ball/the goalie]  caused the outcome? |
|  |  |
| **Event judgements** | |
| Outcome assessment | To what extent do you think your ball scored? |
| Counterfactual thinking | To what extent do you think the ball would have scored if [your ball/the goalie] had gone in the other direction? |
| Causal selection | To what extent did [your ball’s/the goalie’s] moving cause the ball to  [score/not score]? |
| All: Confidence | How confident are you in your response? |
|  |  |
| **Responsibility judgements** | |
| All: Responsibility | How responsible are you for the ball [scoring/not scoring]? |
| How responsible is the goalie for the ball [scoring/not scoring]? |

*Note.* Brackets indicate how each question varies by self- vs. other-focused thinking and/or score vs. miss trials.

**Vividness.** The subjective vividness of mental simulations will be gauged with a 1-5 Likert scale adapted from the second edition of the Vividness of Visual Imagery Questionnaire (VVIQ2; Marks, 1995), with 1 indicating *No image at all, you only “know” that you are thinking of the objects* and 5 indicating *Perfectly clear and lively as real as seeing*. The questions probing vividness will specifically target the mental simulations evoked during retrospective reflections.

**Event judgements.** Participants will next answer a reflection-specific question about the given event using a continuous slider scale, with the leftmost end indicating *Not at all* and the rightmost end indicating *Very much*. These questions will be adapted from those used in Gerstenberg et al., (2017).

**Responsibility judgments.** Lastly, participants will use the same continuous slider scale to ascribe self- and goalie-oriented responsibility for the outcome.

**Behavioral Pilot**

We conducted an online, behavioral pilot study (N = 250) to test the proposed experimental paradigm. We recruited participants from Amazon Mechanical Turk on the basis of [X], All participants provided electronic consent following the procedures approved by the Duke University Internal Review Board.

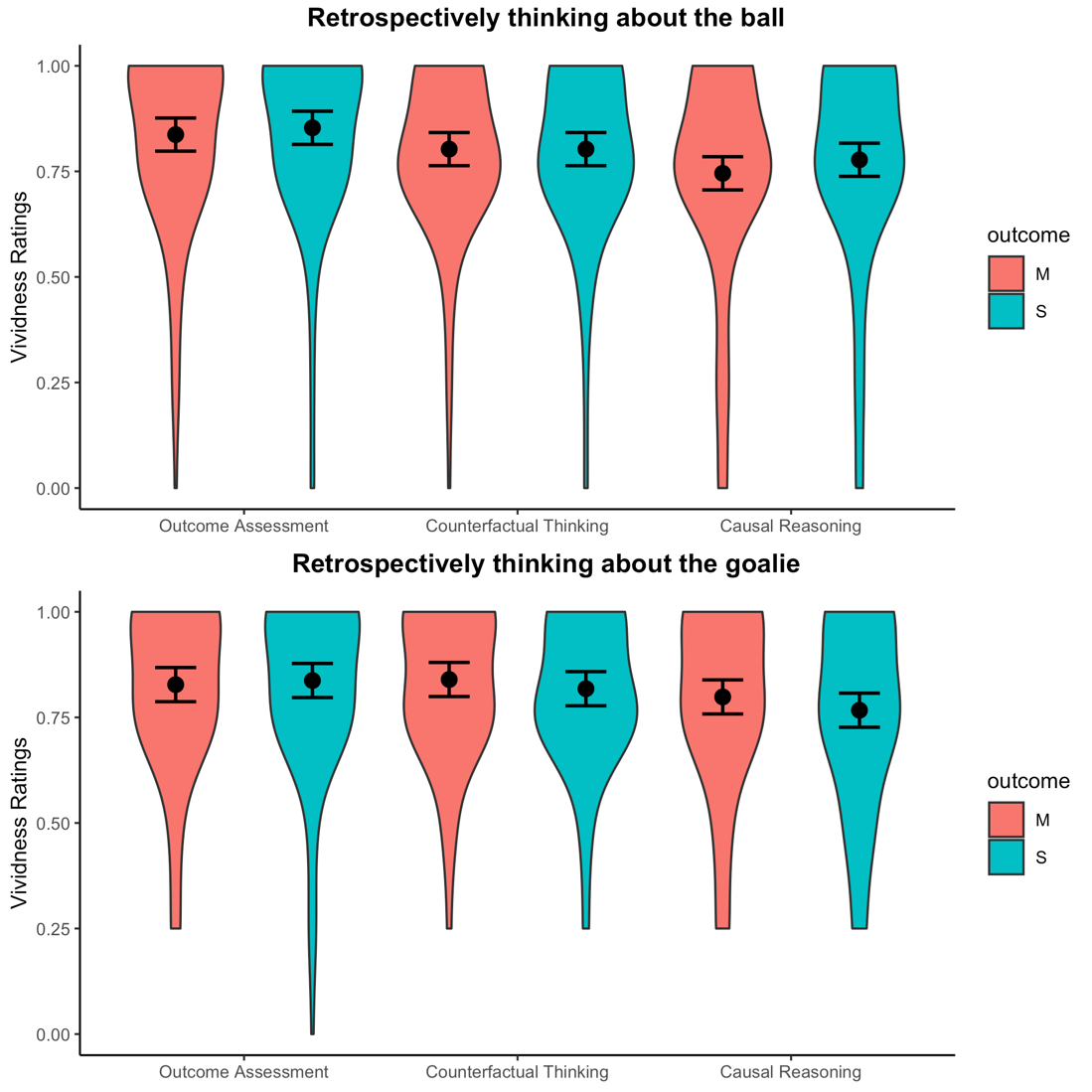
Participants were then randomly assigned to the *ball-focused* or *goalie-focused* conditions. A total of 16 *ball-focused* and 22 *goalie-focused* participants were removed for failing at least one of the two questions used to gauge participants’ engagement and compliance (final N = 212). Participants in this pilot only viewed the upward display orientation and completed 6 total trials (one score and miss trial for each reflection type, randomized).

*Vividness.*

We first assessed whether the vividness of mental simulations varied across retrospective reflections, the success of the outcome, and/or the perspective of the mental simulation. Ratings of vividness (1-5) were normalized on a 0-1 range. Using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015), we then conducted a mixed-effect linear regression analysis that modeled *vividness* as a *thought type* (three levels: outcome assessment [reference group], counterfactual thinking, and causal reasoning) by *perspective* (ball-focused [reference group] and goalie-focused) by *outcome* (two levels: miss [reference group] and score) three-way interaction with *participant* as a random effect.

Findings from the full model are reported in **Table 2.** The reported unstandardized coefficients (B) indicate the predicted change in vividness ratings for each unit increase in the given predictor variable controlling for the other predictor variables. Significance testing was conducted using a two-tailed test with α = .05. As illustrated in **Figure 2**, Tukey-corrected post-hoc comparisons using the *emmeans* package in R (Lenth, 2017)[[1]](#footnote-1) showed that participants reported more vivid mental simulations for outcome assessment than counterfactual thinking (*B* = .04, *SE* = .01, *p* = .004) but only when participants were focused on the ball (personal perspective). Those who were focused on the goalie (impersonal perspective) reported similar vividness ratings for outcome assessment and counterfactual thinking (*B* = .00, *SE* = .02, *p* = .809). Mental simulations for outcome assessment were more vivid than for causal reasoning regardless of whether participants were focused on the ball (*B* = .08, *SE* = .01, *p* < .001) or the goalie (*B* = .05, *SE* = .02, *p* = .001). Further, mental simulations for counterfactual thinking were reported as more vivid than causal reasoning again regardless of whether participants were focused on the ball (*B* = .04, *SE* = .01, *p* = .005) or the goalie (*B* = .05, *SE* = .02, *p* = .002)

These findings indicate that the vividness of mental images evoked during retrospective thinking varied across thought type. This, along with our prediction that the vividness of mental simulation might impact any percept-related event judgments from retrospective thinking, we included vividness ratings as a variable in analyses investigating event ratings (described next).



**Figure 2.** Estimated marginal means and standard errors for normalized ratings indicating the vividness of mental simulations evoked during retrospective outcome assessment, counterfactual thinking, and causal reasoning.

*Table 2.* *Test statistics for all the variables included in the mixed-effect linear regression analysis predicting the vividness of retrospective mental simulations as a retrospective thought type by perspective by outcome three-way interaction*

|  |  |  |  |
| --- | --- | --- | --- |
| ***Predictors*** | ***B*** | ***95% CI*** | ***p*** |
| (Intercept) | **.84** | **.80 – .88** | **< .001** |
| Outcome [score] | .02 | -.03 – .06 | .338 |
| Thought [counterfactual thinking] | -.03 | -.08 – .01 | **.097** |
| Thought [causal reasoning] | **-.09** | **-.13 – -.05** | **< .001** |
| Perspective [goalie] | -.01 | -.07 – .05 | **.741** |
| Outcome [score] \* Thought [counterfactual thinking] | -.02 | -.07 – .04 | .584 |
| Outcome [score] \* imagination [causal reasoning] | .02 | -.04 – .07 | .584 |
| Outcome [score] \* Perspective [goalie] | -.01 | -.07 – .05 | .831 |
| Thought [counterfactual thinking] \* Perspective [goalie] | .05 | -.01 – .11 | .117 |
| Thought [causal reasoning] \* Perspective [goalie] | **-.06** | **.00 – .12** | **.035** |
| (Outcome [score] \* Thought [counterfactual thinking]) \* Perspective [goalie] | -.02 | -.10 – .07 | .712 |
| (Outcome [score] \* Thought [causal reasoning]) \* Perspective [goalie] | -.06 | -.14 – .03 | .173 |
|  |  |  |  |
| ***Random Effects*** |  |  |  |
| σ2 | .02 |  |  |
| τ00 id | .02 |  |  |
| ICC | .46 |  |  |
| N id | 212 |  |  |
| Observations | 1272 |  |  |
| Marginal R2 / Conditional R2 | .02 / .48 |  |  |

*Note.* *B* = unstandardized coefficients; *CI* = confidence interval; Boldface text indicates *p* < .05; id = participant id; Created with the *sjPlot* package in R (Lüdecke, 2020)

*Predicting causal judgments from internal thoughts.*

We next assessed the extent to which participants relied on outcome assessment and counterfactual thinking for causal selection. To that end, we first conducted two linear mixed-effect analyses. *Model 1* modeled event judgements for outcome assessment (*To what extent do you think your ball scored*) as a *perspective* (personal [reference group] and impersonal) by *outcome* (two levels: miss [reference group] and score) by *vividness* three-way interaction with *participant* as a random effect. *Model 2* followed the same design to model event judgments for counterfactual thinking (*To what extent do you think the ball would have scored if [your ball/the goalie] had gone in the other direction*).

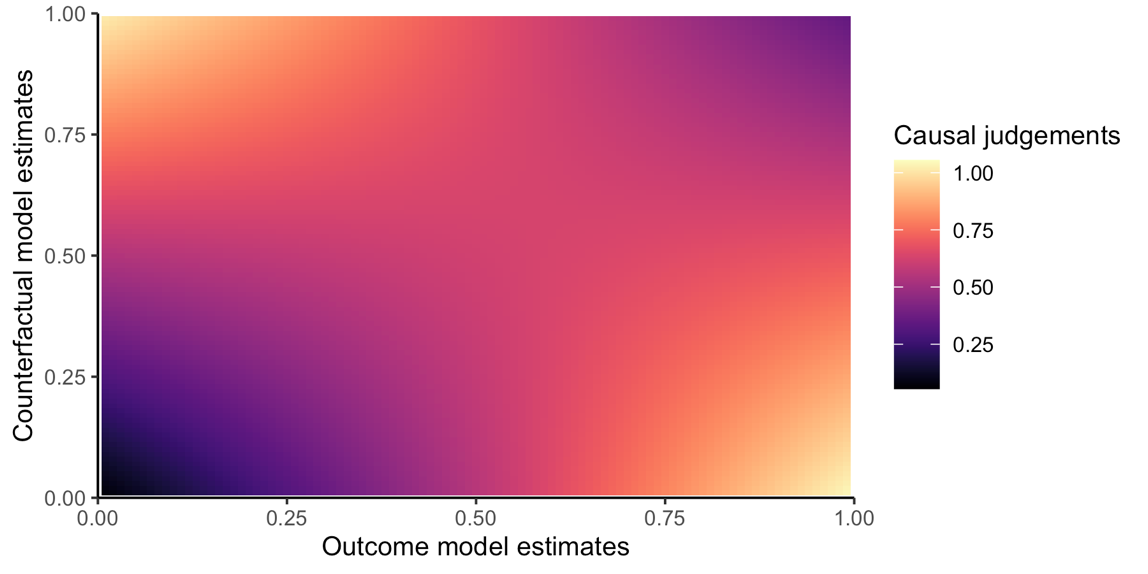
The results from these models are reported in **Table 3**. For both of these models, the reported vividness of retrospective mental simulations significantly predicted event judgements, although this main effect was characterized by a significant outcome by vividness interaction. These findings are further detailed below (*The role of vivid mental simulations*).

We next conducted a mixed-effect linear regression analysis that modeled judgments for causal selection as a *Model 1* (outcome assessment) by *Model 2* (counterfactual thinking) interaction with *participant* as a random effect. These findings showed that Model 1 was a significant predictor of judgments for causal selection (*B* = 1.00, *SE* = .28, *p* < .001) such that larger model estimates corresponded with greater causal judgements. Model 2 was also a significant predictor of causal judgements following the same pattern of results, *B* = .97, *SE* = .29, *p* = .001. These main effects, however, were characterized by a significant Model 1 by Model 2 interaction; *B* = -1.69, *SE* = .39, *p* < .001. As illustrated in **Figure 3**, findings indicated that event judgments for causal selection were greatest under two circumstances. First, they were greatest when Model 1 estimates were high and Model 2 estimates were low. These findings indicate that event judgements for causal selection were predicted by judgements from outcome assessment and suggest that outcome assessment was used for causal selection. Additionally, and critical for adjudicating between process and counterfactual theories, causal judgements were also high when Model 2 estimates were high and Model 1 estimates were low. These findings indicate that event judgements derived from counterfactual thinking also predicted causal judgments, suggesting that participants also engaged, to a certain extent, in counterfactual thinking for causal selection. These findings are, therefore, consistent with counterfactual theories.

*Table 3.* *Test statistics for all variables included in the mixed-effect linear regression analyses predicting event judgements for outcome assessment (Model 1) and counterfactual thinking (Model 2) as a perspective by outcome by vividness three-way interaction*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1: Outcome assessment** | | |  | **Model 2: Counterfactual thinking** | | |
| ***Predictors*** | ***B*** | ***95% CI*** | ***p*** |  | ***B*** | ***95% CI*** | ***p*** |
| (Intercept) | **.48** | **.26 – .70** | **<.001** |  | **.53** | **.28 – .79** | **< .001** |
| Outcome [score] | .26 | -.06 – .58 | .105 |  | -.08 | -.43 – .26 | .640 |
| Perspective [impersonal] | .14 | -.18 – .46 | .400 |  | -.00 | -.39 – .38 | .982 |
| Vividness | **-.26** | **-.51 – -.00** | **.050** |  | **.40** | **.09 – .71** | **.011** |
| Outcome [score] \* Perspective [impersonal] | -.259 | -.72 – .20 | .268 |  | .28 | -.25 – .81 | .295 |
| Outcome [score] \* Vividness | **.43** | **.06 – .80** | **.022** |  | **-.56** | **-.98 – -.14** | **.009** |
| Perspective [impersonal] \* Vividness | -.17 | -.54 – .21 | .384 |  | -.02 | -.47 – .44 | .949 |
| (Outcome [score] \* Perspective [impersonal]) \* Vividness | .31 | -.22 – .84 | .251 |  | -.34 | -.97 – .29 | .290 |
|  |  |  |  |  |  |  |  |
| ***Random Effects*** |  |  |  |  |  |  |  |
| σ2 | .08 |  |  |  | .09 |  |  |
| τ00 | .00 id |  |  |  | .00 id |  |  |
| N | 212 id |  |  |  | 212 id |  |  |
| Observations | 424 |  |  |  | 424 |  |  |
| Marginal R2 / Conditional R2 | .56 / NA |  |  |  | .46 / NA |  |  |

*Note.* *B* = unstandardized coefficients; *CI* = confidence interval; id = participant id; Boldface text indicates *p* < .05



**Figure 3.** Causal judgements as predicted by outcome model estimates and counterfactual model estimates.

*The role of vivid mental simulation*

Both Model 1 (outcome assessment) and Model 2 (counterfactual thinking) showed that vividness was a significant predictor of event judgements, although these main effects were characterized by an outcome by vividness interaction. Specifically, for outcome assessment of miss trials, more vivid mental simulations corresponded with *lower* ratings (*B* = -.34, *SE* = .10, *CI* = -.53 – -.15). That is, reports of more vivid mental simulations corresponded with more extreme judgements of the ball missing the goal. For the outcome assessment of score trials, more vivid mental simulations corresponded with *higher* ratings, or more extreme judgements of the ball scoring (*B* = .25, *SE* = .10, *CI* = .06 – .44). A post-hoc comparison suggested that these effects were significantly different; *B* = -.58, *SE* = .14, *p* < .001. For counterfactual thinking of miss trials, reports of more vivid simulations corresponded with *higher* judgements (*B* = .39, *SE* = .12, *CI* = .16 – .62). That is, more vivid simulations corresponded with more extreme judgements of the ball scoring had it or the goalie moved in the counterfactual direction. Conversely, for score trials, more vivid mental simulations corresponded with *lower* judgements of the ball scoring given the possible counterfactual movements (*B* = -.34, *SE* = .11, *CI* = -.56 – -.12). A post-hoc comparison suggested that these effects were significantly different; *B* = .73, *SE* = .16, *p* < .001. Considering all of these results collectively, then, reported vividness corresponded with more extreme judgements of the actual and counterfactual outcome, as hypothesized.

Because these results show that the vividness of mental simulations predict event judgments, and that participants seemed to engage in both outcome assessment and counterfactual thinking for causal selection, it should follow that vividness would also significantly predict causal judgements. To test this idea, we conducted another mixed-effect linear regression analysis modeled causal judgements (*To what extent did [your ball’s/the goalie’s] moving cause the ball to [score/not score]*) as a *perspective* (personal [reference group] and impersonal) by *outcome* (two levels: miss [reference group] and score) by *vividness* three-way interaction with *participant* as a random effect. These results are reported in **Table 4** and indeed showed that the vividness of the mental simulation significantly predicted causal judgements. Specifically, more vivid mental simulations corresponded with higher causal judgements. Implications for these findings are further described next.

*Table 4. Test statistics for all variables included in the mixed-effect linear regression analysis predicting causal judgements as a perspective by outcome by vividness three-way interaction*

|  |  |  |  |
| --- | --- | --- | --- |
| ***Predictors*** | ***B*** | ***95% CI*** | ***p*** |
| (Intercept) | .57 | .44 – .70 | **<.001** |
| Outcome [score] | .12 | -.04 – .28 | .144 |
| Perspective [goalie] | -.03 | -.24 – .18 | .784 |
| Vividness | .24 | .07 – .40 | **.004** |
| Outcome [score] \* Condition [goalie] | -.16 | -.41 – .09 | .219 |
| Outcome [score] \* Vividness | -.06 | -.26 – .14 | .560 |
| Perspective [goalie] \* Vividness | .07 | -.18 – .33 | .574 |
| (Outcome [score] \* Perspective [goalie]) \* Vividness | .04 | -.27 – .35 | .799 |
|  |  |  |  |
| ***Random Effects*** |  |  |  |
| σ2 | .03 |  |  |
| τ00 id | .02 |  |  |
| ICC | .38 |  |  |
| N id | 212 |  |  |
| Observations | 424 |  |  |
| Marginal R2 / Conditional R2 | .09 / .44 |  |  |

*Note.* *B* = unstandardized coefficients; *CI* = confidence interval; id = participant id; Boldface text indicates *p* < .05

*Discussion*

[I think we should have a paragraph summarizing the behavioral pilot?]

**Proposed analyses and anticipated results**

*Vividness across thought types*

Following our hypotheses and pilot results, it is possible that the vividness of mental simulations will vary across the type of retrospective thought, the success of the outcome, and/or the perspective of the mental simulation. To test this idea, we will normalize ratings of vividness and conduct the same mixed-effect linear regression analysis as in the pilot. That is, we will model *vividness* as a *thought* (three levels: outcome assessment [reference group], counterfactual thinking, and causal reasoning) by *perspective* (personal [reference group] and impersonal) by *outcome* (two levels: miss [reference group] and score) three-way interaction with *participant* as a random effect.

We anticipate replicating our pilot results such that the vividness of mental simulations for outcome assessment will be reported as greater than those for counterfactual thinking and causal selection. This is because X. It is possible that, as observed in the pilot, these findings will vary across perspective given that X. We do not anticipate a main effect of outcome. Especially if the vividness of mental simulations varies across retrospective thought, we will include vividness ratings as a covariate in all other analyses.

*Predicting causal judgments*

We will use the same modeling strategies to investigate

Kristina, I was thinking of using either a generalized additive model (GAM), or a Gaussian Process (GP) model for the eye movements. Of course, though, on a first pass we’ll want to look at heatmaps of the fixations  to see if there’s any noticeable difference between conditions before running the models.

**General Discussion**

**References**

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*(2), 365.

Clarke, R., Shepherd, J., Stigall, J., Waller, R. R., & Zarpentine, C. (2015). Causation, norms, and omissions: A study of causal judgments. *Philosophical Psychology*, *28*(2), 279-293.

Damiano, C., & Walther, D. B. (2019). Distinct roles of eye movements during memory encoding and retrieval. *Cognition*, *184*, 119-129.

Dowe, P. (2000). Causality and explanation. *The British journal for the philosophy of science*, *51*(1), 165-174.

Goldvarg, E., & Johnson‐Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive science*, *25*(4), 565-610.

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

Hannula, D. E., Ryan, J. D., Tranel, D., & Cohen, N. J. (2007). Rapid onset relational memory effects are evident in eye movement behavior, but not in hippocampal amnesia. *Journal of cognitive neuroscience*, *19*(10), 1690-1705.

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.

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

Knobe, J. (2009). Folk judgments of causation. *Studies in History and Philosophy of Science Part A*, *40*(2), 238-242.

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

Morris, A., Phillips, J. S., Icard, T., Knobe, J., Gerstenberg, T., & Cushman, F. (2018). Causal judgments approximate the effectiveness of future interventions.

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, *26*(2), 207-231.

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

Lüdecke D (2020). *sjPlot: Data Visualization for Statistics in Social Science*. doi: [10.5281/zenodo.1308157](https://doi.org/10.5281/zenodo.1308157), R package version 2.8.3, <https://CRAN.R-project.org/package=sjPlot>.

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

Salmon, W. C. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science*, *64*(3), 461-477.

Walsh, C. R., & Sloman, S. A. (2011). The meaning of cause and prevent: The role of causal mechanism. *Mind & Language*, *26*(1), 21-52.

Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans: Factors affecting judgments of response-outcome contingencies under free-operant procedures. *Learning and motivation*, *14*(4), 406-432.

Wolff, P. (2007). Representing causation. *Journal of experimental psychology: General*, *136*(1), 82-111.

1. Comparisons were adjusted using the Tukey method for multiple comparisons. [↑](#footnote-ref-1)