**Stage 1 Registered Report: Eye tracking mental simulation during retrospective causal reasoning**

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**

Successfully navigating everyday tasks requires people to evaluate cause-and-effect relationships. Such causal reasoning enables people to judge whether something was causally relevant to the outcome and modify behavior accordingly to achieve optimal outcomes. How people form causal judgments, however, remains unclear, given that these judgments are influenced by a variety of factors, such as perceived norms and expectations (e.g., Knobe, 2009; Knobe & Fraser, 2008), the degree to which potential causes and effect 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). From these findings, there are conflicting theories regarding how people manage to form causal judgments.

As one prominent theory of causal reasoning, *process theories* argue that people will focus on the direct interaction between a covarying event and outcome, assessing whether there was an exchange of some conserved quantity, such as energy or momentum (Dowe, 2000; Salmon, 1997; Wolff, 2007). Supportive evidence shows that people tend to judge something as more causally relevant 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). These actions are usually judged are more causally relevant when they directly precede the outcome, and possible enabling conditions (e.g., effective coaching strategies) are often judged as less causally relevant (e.g., Goldvarg & Johnson-Laird, 2001). Therefore, process theories predict that people preferentially consider what *actually happened* between the potential causal event and the outcome to form a causal judgment.

As one alternative account, *counterfactual theories* argue that people compare what actually happened with what reasonably might have happened had the potential causal event been absent or altered in some way (e.g., Lewis, 1973). This form of counterfactual thinking enables people to consider alternative scenarios and outcomes to assess whether the potential causal event made a difference in the actual outcome. Counterfactual thinking can be highly episodic, with memory-like spatiotemporal mental simulations that resemble, but are altered from, the actual remembered event (Roese & Epstude, 2017; Byrne 2016). Therefore, counterfactual thinking 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). Counterfactual thinking frequently occurs when outcomes are negative and/or unexpected (Roese, 1997), but 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.

The fundamental difference between process and counterfactual theories, thus, is the extent to which people consider what *actually* happened with what *counterfactually* could have happened when forming a causal judgment. Toward adjudicating between these prominent theories, Gerstenberg et al., (2017) tracked participants’ eye movements to investigate what visuospatial information—such as about the form, structure, and spatial relationship of objects—people fixated in the process of forming a causal judgment. Eye movements provide a spatially precise measure of the overt allocation of attention. This is because visual inputs are only acquired during fixations—periods of time when the eye remain relatively still—but are physically [ref] and cognitively blocked [ref] during saccades—the ballistic eye movements that relocate fixations. Accordingly, Gerstenberg et al. (2017) hypothesized that how people fixated on visuospatial information would reveal how such information was evaluated while forming a causal judgment.

Gerstenberg et al., (2017) asked participants to watch videos of two balls (Balls A and B) moving toward a goal and colliding into each other. While watching the videos, participants, as between-subject conditions, thought about 1) the extent to which Ball B scored (outcome assessment), the extent to which it would have scored had Ball A not been present (counterfactual thinking), or whether Ball A caused Ball B to score or not score (causal judgment). The findings showed that participants who engaged in counterfactual thinking and causal judgment—compared to those who engaged in outcome assessment—had a greater tendency to fixate in the spatial locations where Ball B would have moved had Ball A not interfered (the *counterfactual movement* of Ball B). The tendency to fixate on counterfactual movements corresponded with Ball A being judged as more causally relevant to the outcome of Ball B. The authors therefore inferred that participants tended to simulate the counterfactual movements of Ball B to judge the causal relevance of Ball A.

Unfortunately, this conclusion is unwarranted given a critical limitation of the study, namely that Ball B’s counterfactual movement was spatially equivalent to its original path before Ball A’s interference. This made it impossible to disambiguate whether fixations to this location reflected counterfactual simulations or instead future-directed hypothetical simulations of Ball B’s original projected movement (e.g., extrapoloation; Howe & Holcombe, 2012; Luu & Howe, 2015). The authors acknowledged this limitation but suggested further that fixations toward Ball B’s projected movement would still allow participants to acquire the visuospatial information needed for generating counterfactual contrasts later once the visual input was removed. This account, though, was not tested, given that the recorded eye movements did not reflect whether and how participants *retrospectively* reflected on what could have occurred to form the causal judgment.

**The current study**

The current research will address the limitation of Gerstenberg et al. (2017) and further adjudicate between process and counterfactual theories by investigating how people overtly attend to visuospatial information both while it is perceptually available as well as while people retrospectively reflect on the event to form a causal judgment. Participants will complete a ball-shooting game wherein they will, as randomized between-subject conditions, shoot a ball toward a goal trying to score (*ball-focused*), or they will move a goalie trying to block the ball from scoring (*goalie-focused*). Participants will first decide how to move the ball or goalie (left or right), then they will watch a video of the event and outcome (e.g., the ball moving toward the goal, and the goalie blocking the ball from scoring). People will then, as randomized within-subject conditions, retrospectively 1) reflect on the actual outcome (outcome assessment), 2) consider a counterfactual event and outcome (counterfactual thinking), or 3) form a causal judgment about what caused the ball to score or not score (causal judgment). Eye movements will be record both while participants watch the video and while they retrospectively reflect on the outcome to assess what visuospatial information is attended during encoding and mentally simulated later during reflection.

During encoding, participants will be unaware of how they will need to later reflect on the outcome. This will require participants to prepare for any type of reflection by sampling all potentially relevant visuospatial information. As such, eye movements during initial encoding are not predicted to vary across trials and will thus serve as a baseline understanding of how visuospatial information is encoded without any specific top-down, goal-direct impacts. It is likely that participants will frequently fixate on the moving objects because task-relevant motion is a highly salient visual input, likely to capture attention (e.g., Hillstrom & Yantis, 1994; Yantis & Egeth, 1999). Further, this visuospatial information is relevant to assessing the actual event and outcome. That said, consistent with the conclusions from Gerstenberg et al., 2017, people may also sometimes fixate on counterfactual movements, which would facilitate the counterfactual simulations that may need to be generated later during reflection.

While participants reflect on the outcome, they will look at a blank screen. Past research consistently shows that when people look at a blank screen and construct a mental simulation of previously viewed visuospatial information, they tend to systematically move their eyes to the spatial locations once occupied by the original visual information (e.g., Altmann, 2004; Brandt & Stark, 1997; Richardson & Spivey, 2000). This tendency is correlated with better recall of the visual information, the self-reported vividness of the mental image, and with similarities in brain activation patterns between initial encoding and later imagery (e.g., Bone et al., 2019; Gurtner, Bishof, & Mast, 2019; Laeng & Teodorescu, 2002; Martarelli & Mast, 2013). 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 eye movements can delineate how visuospatial information is mentally simulated, suggesting the utility of using eye movements to investigate the mental simulations created in the process of forming causal judgments.

Process and counterfactual theories offer unique predictions about how people will allocate eye movements to mentally simulated visuospatial information while forming causal judgements. Process theories predict that people will only look at mentally simulated visuospatial information related to the *actual* event because this information is sufficient for forming casual judgments. Counterfactual theories, however, predict that people will additionallylook at visuospatial information relevant to the *counterfactual* movements of the objects. Doing so would enable contrasts between actual and counterfactual events to be generated and considered for forming causal judgments. The critical difference in these predictions, therefore, is the extent to which people will fixate on mentally simulated visuospatial information related to the actual or counterfactual event and outcome. The current work will, therefore, model participants’ eye movements as they retrospectively reflect on the *actual* event and outcome (outcome assessment) and a *counterfactual* event and outcome (counterfactual thinking). These modeled eye movements will then be used to model the eye movements evoked while participants form a causal judgment. The findings will reveal the extent to which people fixate on visuospatial information related to the actual or counterfactual event and outcome.

**Methods**

**Participants**

Our data will be analyzed within a mixed-effects framework for which there is not agreed upon method for properly conducting an a priori power analysis to estimate a target sample size (ref). To circumvent this challenge, we will first establish a target sample size of 40 participants, which was the sample size used in Gerstenberg et al. (2017). There will be a total of 44 participants recruited to allow for possible attrition, such as due to cancellations or technical issues. We will then use Bayesian analytic approaches for all our analyses (see Proposed Analyses).

Participants will be volunteers between the ages of 18-35 years old. They will be recruited from Duke University and the local community. Informed consent will be obtained from each participant following procedures approved by the University Institutional Review Board, and participants will be compensated $12/hr for participating in the laboratory.

**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 1A**) 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) an orange, circular 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. Half of the participants will see the upward orientation for the first half of trials, while the other half of participants will see the downward orientation first.

This display configuration boasts several empirical advantages to adjudicating between process andcounterfactual theories of causal selection. First, the actual movement of each object and its only corresponding counterfactual movement reside in opposite visual hemifields, which disambiguates the projected movement of each object from their corresponding counterfactual movement. For example, as illustrated in **Figure 1A,** if the ball moves diagonally to the left, its only possible counterfactual movement would be a diagonal trajectory to the right. Any anticipatory, future-oriented hypothetical simulations of each object’s projected movement will likewise reside in opposite visual hemifields of the counterfactual movement. This display configuration thus addresses the critical limitation of Gerstenberg et al. (2017).

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 as well as halfway through to study. A one-point calibration will be used before each trial 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, participants will be randomly assigned to either the *ball-focused* or the *goalie-focused* conditions for equally sized groups*.* These conditions will be defined by what object participants will chose how to move, attend, and mentally simulate. The ball-focused participants will be instructed to *Shoot the ball and score as many times as possible*, whereas the goalie-focused participants will be instructed to *Defend the goal and block the ball as many times as possible*. Participants will then receive detailed task instructions and watch several instructional videos to learn the starting position of each object, the speed and angle by which the objects move, and how they may interact.

After following the 9-point calibration procedures, participants will begin the experiment. Participants will start each trial by first deciding whether to move their object to the left or right (**Figure 1B**, **Decision**). They will then watch a video of the outcome to encode the outcome, namely whether the ball scored or the goalie blocked the ball from scoring (**Figure 1B**, **Encoding**). 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 a centrally presented prompt for two seconds, either *Remember*, *What If*, or *Cause* (**Figure 1B**, **Prompt**). This prompt will tell participants how they should next reflect on the past event and outcome while looking at a blank screen (**Figure 1B**, **Reflection**). 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 judge the casual relevance of the ball or goalie’s movement in the outcome (causal judgment).

During reflection, participants will be asked to create a vivid mental simulation of the object of focus. This instruction will be emphasized with, *As you do so, try your best to visualize it in your mind. Try to create a vivid mental image.* Past research has suggested that peoples’ ability to create a vivid visual mental simulation can range considerably (ref). To address this challenge, participants will be told, *This can sometimes be difficult or come very naturally. Either is okay!*. Then, along with other end-of-trial questions, participants will report the vividness of their mental simulation using an unnumbered, continuous slider scale with anchors adapted from the second edition of the Vividness of Visual Imagery Questionnaire (VVIQ2; Marks, 1995). The leftmost end will indicate *No image at all, you only “know” that you are thinking of the objects*, and the right most end will indicate *Perfectly clear and lively as real as seeing*. The anchors in between will include, “*Dim and vague; flat*,” “*Moderately clear and lively*,” and “*Clear and lively*” in sequential order.



**Figure 1.** A) Example video display orientations. In the exampled *upward orientation*, the ball and the goalie are projected to move to the right, and their counterfactual movements would be to the left. This trial would result in a *miss* because the ball and the goalie are projected to move in the same direction. In the exampled *downward orientation*, the ball is projected to move to the left and its counterfactual movement would be to the right. The goalie is projected to move to the right, and its counterfactual movement would be to the left. This trial would result in a *score* because the ball and the goalie are projected to move in opposite directions. B) Example trial sequence illustrating an *outcome assessment* trial. Participants will first decide whether to shoot the ball to the left or to the right of the goal (Decision). They will then watch a video of the outcome to encode whether the ball scored or was blocked by the goalie (Encoding). A word will then appear in the center of the screen indicating how participants should next retrospectively reflect on the outcome (Prompt). While looking at a blank screen (Reflection), participants will then retrospectively reflect on what just occurred (*outcome assessment*), a possible alternative outcome (*counterfactual thinking*), or the candidate cause of the outcome (*causal judgment*). Last, participants will answer reflection-specific questions about the given outcome (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 judgment trials will inquire about the causal relevance of either the ball or the goalie in the outcome.

After reflection and reporting the vividness of the mental simulation, participants will answer the condition-specific event-related questions (**Figure 1B, Judgments**), with are listed in **Table 1**. Participants will use an unnumbered, continuous slider scale, with the leftmost end indicating *Not at all*, the rightmost end indicating *Very much*, and the center indicating *Moderately*. Participants will use a similar slider scale to indicate the degree of confidence they have in their response.

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

|  |  |
| --- | --- |
| **Vividness of mental simulation** | |
| Outcome assessment | How vividly could you visualize how [the ball/the goalie] just moved? |
| Counterfactual thinking | How vividly could you visualize what would have happened if [the  ball/the goalie] had moved in the other direction? |
| Causal selection | How vividly could you visualize whether [the ball/the goalie]’s movement  caused the outcome? |
|  |  |
| **Condition-specific event judgements** | |
| Outcome assessment | To what extent do you think that the ball scored? |
| Counterfactual thinking | To what extent do you think the ball would have scored if [the ball/the goalie] had moved [left/right]? |
| Causal selection | To what extent did [the ball/the goalie]’s moving [left/right] caused the ball to  [score/not score]? |
| All: Confidence | How confident are you in your response? |

*Note.* Brackets indicate how each question varies by ball- vs. goalie-focused conditions as well as score vs. miss trials.

The next trial will start once all questions are answered. Participants will complete 4 blocks of 12 trials, for a total of 48 trials. At the end of each block, participants will be reminded of the reflection-specific instructions. All experiment procedures are estimated to take no more than 60 minutes, which includes the initial calibration of the eye tracker. Our study as a whole will be structured as a 2 (reflected object: ball-focused or goalie-focused) x 2 (display layout: upward or downward) x 2 (mental construct: encoding or reflection) x 3 (reflect type: outcome assessment, counterfactual thinking, or causal selection) x 2 (outcome: score or miss) design.

**Online pilot**

To assess the efficacy of the proposed paradigm, we conducted an online pilot study without eye tracking. Participants (*N* = 44) were recruited through the Prolific online research participant pool on the basis that they were located in the United States and had minimal of 95% approval rating. All participants provided electronic consent following the procedures approved by the Duke University Internal Review Board and were compensated $9 for participating.

Participants were randomly assigned to the *ball-focused* or *goalie-focused* conditions. There were X participants from the ball-focused and X participants from the goalie-focused conditions that were removed for failing at least one of the check questions. These exclusionary criteria resulted in a final sample size of *N* = 35 (*n* = 14 ball-focused; *n* = 21 goalie-focused). Individual trials were excluded if participants could not accurately remember the direction that the object [opposite of what they were supposed to focus on] moved to screen trials where participants may not have been paying close attention.

The analyses focused on modeling response judgements for the different simulation types. First, we separately modeled normalized response judgements for outcome assessment and counterfactual thinking. Specifically, using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015), we used a mixed-effect linear regression analysis to model response judgments for outcome assessment as a *reflected object* (two levels: ball-focused and goalie-focused) by *outcome* (two levels: miss and score) by *vividness* (normalized, continuous) three-way interaction with random slopes and random intercepts for outcome for each *participant*. We conducted the same regression analysis to model response judgments for counterfactual thinking.

The results from these models are illustrated in **Figure 2** and reported in **Table 2**. The outcome assessment model showed a significant effect of outcome, which showed greater judgment estimates for score trials than miss trials to the question *To what extent do you think that the ball scored?.* This main effect was characterized by a *reflected object* by *outcome* by *vividness* three-way interaction. Using the *emmeans* package in R (ref), pairwise comparisons showed that ball-focused participants showed a positive *outcome* by *vividness* trend for score trials (*β* = .53, 95% *CI* = .25 – .82), which indicates that greater vividness judgments were associated with greater judgments that the ball actually scored. When the ball missed, ball-focused participants showed a negative *outcome* by *vividness* trend (*β* = -.03, 95% *CI* = -.31 – .25), which indicated that greater vividness judgments tended to correspond with greater judgments that the ball did not score. The contrast between these trends was significant, *β* = -.56, *SE* = .22, *t*(411) = -2.61, *p* = .009. Goalie-focused participants showed no difference in the *outcome* by *vividness* trend for score (*β* = .10, 95% *CI* = -.02 – .22) and miss (*β* = .04, 95% *CI* = -.14 – .22) trials, *β* = -.06, *SE* = .12, *t*(442) = -0.49, *p* = .623. Together, these findings indicate that participants correctly judged whether the ball scored or missed, and the judgments from the ball-focused participants were estimated to be more extreme with reportedly more vivid mental simulations.

The counterfactual thinking model also showed a significant effect of outcome, which showed greater judgment estimates for miss trials than score trials to the question *To what extent do you think the ball would have scored if [the ball/the goalie] had moved [left/right]?*. Although there was no significant three-way interaction, the effect of outcome was characterized by a significant *outcome* by *vividness* interaction and a trending *reflected object* by *outcome* interaction. A pairwise comparison showed that ball-focused participants showed no difference in the *outcome* by *vividness* trend for score (*β* = .07, 95% *CI* = -.19 – .33) and miss (*β* = .20, 95% *CI* = -.12 – .53) trials, *β* = .13, *SE* = .22, *t*(287) = 0.60, *p* = .552. Goalie-focused participants, however, showed a negative *outcome* by *vividness* trend for score trials (*β* = -.28, 95% *CI* = -.48 – -.07) but a positive trend for miss trials (*β* = .18, 95% *CI* = .02 – .34) trials, and these trends were significantly different, *β* = .46, *SE* = .14, *t*(295) = 3.18, *p* = .002. These findings collectively indicate that participants could correctly identify the counterfactual outcome when the ball actually scored or missed, and that the judgments from the goalie-focused participants were estimated to be more extreme with reportedly more vivid mental simulations.

The collective findings from the outcome assessment model and the counterfactual thinking model showed the efficacy of our empirical approach. First, participants correctly ascertained the actual and counterfactual outcomes of the ball scoring or not scoring. Further, the reported vividness of the mental simulation—which served as a proxy for whether participants might generate eye movements during mental simulation—sometimes corresponded with more extreme actual and counterfactual outcome judgements.

Chart

Description automatically generated

**Figure 2.** Model estimates (black solid line) and 95% CI (color shaded regions) for the outcome assessment model and counterfactual thinking model; individual data estimates are represented by the colored dots. The outcome assessment model showed a significant effect of outcome that was characterized by a *reflected object* by *outcome* by *vividness* three-way interaction. The counterfactual thinking model showed a significant effect of outcome that was characterized by a significant *outcome* by *vividness* interaction and a trending reflected *object* by *vividness* interaction.

*Table 3.* *Test statistics for each variable included in the mixed-effect linear regression analyses modeling response judgements for outcome assessment and counterfactual thinking*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Outcome assessment** | | | |  |  | **Counterfactual thinking** | | | |  |
| ***Predictors*** | ***B*** | ***95% CI*** | | ***p*** |  |  | ***B*** | ***95% CI*** | | ***p*** |  |
| (Intercept) | .53 | .50 – .57 | | **< .001** | \* |  | .51 | .48 – .55 | | **< .001** | **\*** |
| Reflected object | -.23 | -.06 – .01 | | .217 |  |  | .00 | -.03 – .04 | | .971 |  |
| Outcome | -.39 | -.45 – -.33 | | **< .001** | \* |  | .40 | -.34 – .45 | | **< .001** | \* |
| Vividness | .04 | .01 – .06 | | **.002** | \* |  | .01 | -.02 – .04 | | **.441** |  |
| Reflected object \* Outcome | -.02 | -.08 – .05 | | **.624** |  |  | .01 | -.04 – .07 | | **.685** |  |
| Reflected object \* Vividness | .02 | -.00 – .04 | | **.080** |  |  | .02 | -.00 – .05 | | **.102** |  |
| Outcome [score] \* Vividness | -.04 | -.06 – -.01 | | .011 | \* |  | .04 | .01 – .07 | | .023 | \* |
| (Reflected object \* Outcome) \* Vividness | -.03 | -.06 – -.00 | | **.039** | \* |  | -.02 | -.05 – .01 | | **.207** |  |
|  |  |  | |  |  |  |  |  | |  |  |
| ***Random Effects*** |  |  | |  |  |  |  |  | |  |  |
| σ2 | .02 |  | |  |  |  | .03 |  | |  |  |
| τ00 | .01 ID |  | |  |  |  | .01 ID |  | |  |  |
| τ11 | .03 ID.Outcome | |  | |  |  | .02 ID.Outcome | |  | |  |
| *ρ*01 | .91 ID |  | |  |  |  | -.63 ID |  | |  |  |
| ICC | .64 |  | |  |  |  | .55 |  | |  |  |
| N | 35 ID |  | |  |  |  | 35 ID |  | |  |  |
| Observations | 501 |  | |  |  |  | 483 |  | |  |  |
| Marginal R2 / Conditional R2 | .71 / .90 |  | |  |  |  | .74 / .88 |  | |  |  |

*Note.* *B* = unstandardized coefficients; *CI* = confidence interval; \* = *p* < .05; ID = participant ID; σ2 = within-group variance; τ = between-group variance; *ρ* = population correlation coefficient; ICC = intraclass correlation coefficient

We next used the outcome assessment and counterfactual thinking model estimates to model normalized causal judgments. Both process and counterfactual theories of causation predict that people use some degree of outcome assessment to derive causal judgments, which suggests that the outcome assessment model here would predict causal judgments. Counterfactual theories further predict that people will, to a certain extent, rely on counterfactual thinking to derive causal judgments, suggesting that the counterfactual thinking model here might also predict causal judgments. We further considered a possibility that ball-focused and goalie-focused participants would differentially rely on outcome assessment and counterfactual thinking to derive casual judgments, such as, for one example, if ball-focused participants only relied on actual movements of the ball but goalie-focused participants relied on both the actual and counterfactual movements of the goalie. Therefore, we conducted a mixed-effects linear regression analysis modeling causal judgment as a *reflected object* (two levels: ball-focused and goalie-focused) by *outcome assessment model* by *counterfactual thinking model* three-way interaction with random slopes for each *participant*.

The findings are illustrated in **Figure 3** and reported in **Table 3**. Both the *outcome assessment model* and *counterfactual thinking model* predicted causal judgments, but these effects were characterized by a significant interaction. Specifically, causal judgments were greater when the outcome assessment model estimates were greater (i.e., higher estimates that the ball actually scored) and the counterfactual thinking model estimates were lower (i.e., lower estimates that the ball counterfactually would have scored). Causal judgments were also greater when outcome assessment model estimates were lower (i.e., lower estimates that the ball actually scored) and counterfactual thinking model estimates were greater (i.e., greater estimates that the ball counterfactually would have scored). Collectively, these findings suggest that the outcome assessment and counterfactual thinking models both predicted causal judgments, which is consistent with the idea that participants did, to some extent, rely on counterfactual thinking for deriving causal judgments.

The effects of *outcome assessment model* and *counterfactual thinking model* were each also characterized by a significant interaction with *reflected object*. These interactions further show that […]

Chart, scatter chart

Description automatically generated

**Figure 3.** Causal selection judgements as predicted by outcome assessment model estimates and counterfactual thinking model estimates; black dots indicate individual data estimates. Causal selection judgments were greatest when outcome assessment model and counterfactual thinking model estimates were high, suggesting that participants relied on both outcome assessment and counterfactual thinking for causal selection.

Collectively, these findings […]

*Table 3.* *Test statistics for each variable included in the mixed-effect linear regression analyses modeling causal selection judgements*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Predictors*** | ***B*** | ***95% CI*** | ***p*** |  |
| (Intercept) | .79 | .73 – .86 | < .001 | \* |
| Reflected object | -.04 | -.10 – .03 | .299 |  |
| Outcome assessment model | .09 | .03 – .15 | .002 | \* |
| Counterfactual thinking model | .08 | .03 – .14 | .005 | \* |
| Reflected object \* Outcome assessment model | .10 | .04 – .16 | **.001** | \* |
| Reflected object \* Counterfactual thinking model | .09 | .03 – .15 | **.002** | \* |
| Outcome assessment model \* Counterfactual thinking model | -.15 | -.21 – -.09 | < .001 | \* |
| (Reflected object \* Outcome assessment model) \* Counterfactual thinking model | -.04 | -.10 – .02 | **.228** |  |
|  |  |  |  |  |
| ***Random Effects*** |  |  |  |  |
| σ2 | .02 |  |  |  |
| τ00 | .01 ID |  |  |  |
| ICC | .39 |  |  |  |
| N | 35 ID |  |  |  |
| Observations | 497 |  |  |  |
| Marginal R2 / Conditional R2 | .13 / .47 |  |  |  |

*Note.* *B* = unstandardized coefficients; *CI* = confidence interval; \* = *p* < .05; ID = participant ID; σ2 = within-group variance; τ = between-group variance; ICC = intraclass correlation coefficient

**Proposed eye tracking analyses**

The eye movements will be modeled using multivariate Gaussian process (GP) models—a probabilistic, non-parametric modelling method that can estimate patterns in nonlinear data. This Bayesian approach comprehensively models prediction uncertainty and can describe the expected inter- and intra-spatiotemporal correlations in eye movements [18]. Therefore, multivariate GP models will afford critical advantages to modeling eye movements across different mental constructs (i.e., encoding and simulation) and reflection types (i.e., outcome assessment, counterfactual thinking, and causal judgment).

Using two separate multivariate GP regression models for outcome assessment and counterfactual thinking, we will first model eye movements (location and duration of fixations as a single fixation map) as a *reflected object* (two levels: ball-focused and goalie-focused) by *mental construct* (two levels: encoding and simulation) by *outcome* (two levels: miss and score) three-way interaction with random slopes for *participant*. Then, using the predictions from these two models, we will model eye movements for the causal judgment trials as a *reflected object* by *outcome assessment model* by *counterfactual thinking model* three-way interaction with random slopes for *participant*.

**Predicted results**

**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.