Does motivation influence causal judgments via biased counterfactual simulation?

Ross Kempner

YC Leong

Tobias Gerstenberg

Own Goal or No?

left side:

- +$3 for accuracy
- +$7 if own goal
- -$7 if not
- own goal :)

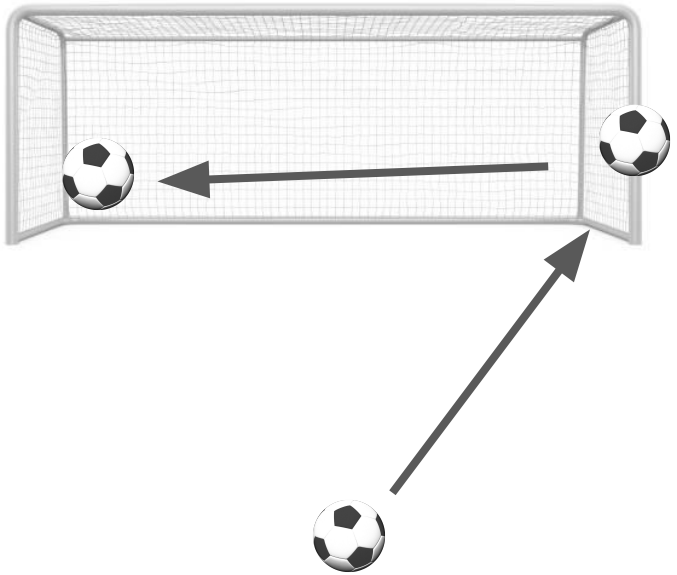
right side:

- +$3 for accuracy
- - $7 if own goal
- +$7 if not
- no own goal :)

Simulated Counterfactual Worlds, Remove the Goalie

left side = no goal

right side = goal



Agenda:

- ➔ Motivated Reasoning
  - ◆ in causal inferences
  - ◆ motivated perception
- ➔ Counterfactual Simulation Model (CSM)
- ➔ Stimuli/Video Clips
- ➔ Experiment 1: Mturk
  - ◆ Hypothesis
  - ◆ Design
  - ◆ Results
- ➔ Experiment 2: Eye-tracking
  - ◆ Design
- ➔ Future Directions

Motivated Reasoning: using a biased set of cognitive processes to arrive at a desired conclusion

The Case for Motivated Reasoning

Ziva Kunda
Princeton University

It is proposed that motivation may affect reasoning through reliance on a biased set of cognitive processes—that is, strategies for accessing, constructing, and evaluating beliefs. The motivation to be accurate enhances use of those beliefs and strategies that are considered most appropriate, whereas the motivation to arrive at particular conclusions enhances use of those that are considered most likely to yield the desired conclusion. There is considerable evidence that people are more likely to arrive at conclusions that they want to arrive at, but their ability to do so is constrained by their ability to construct seemingly reasonable justifications for these conclusions. These ideas can account for a wide variety of research concerned with motivated reasoning.

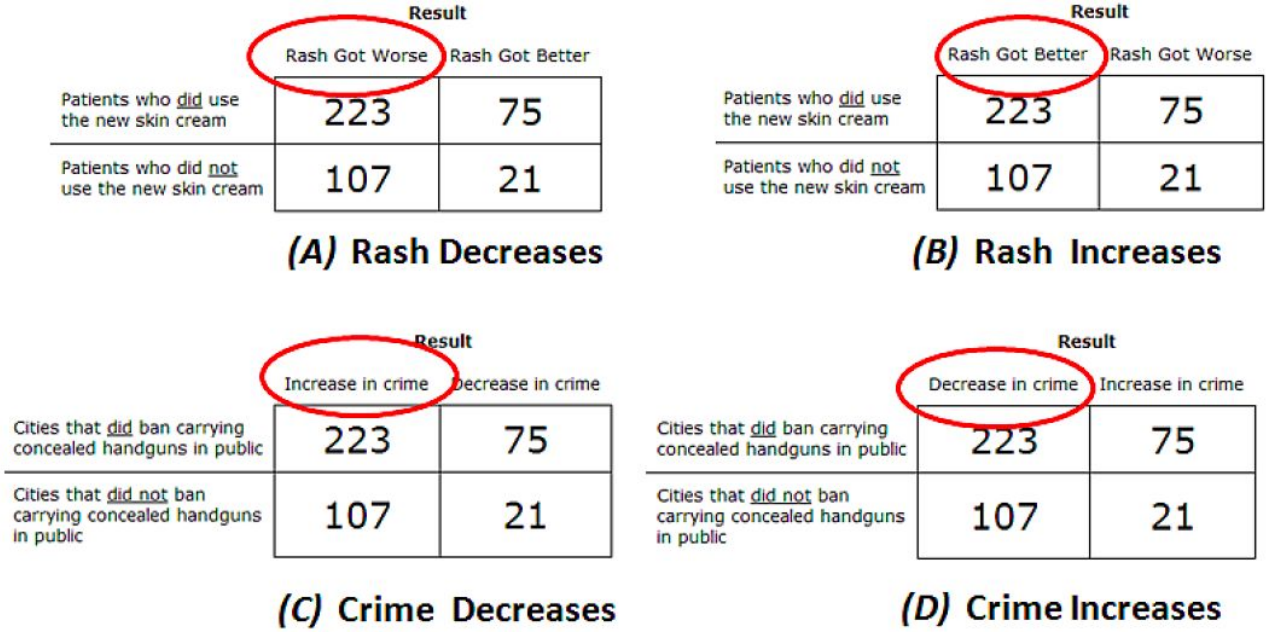
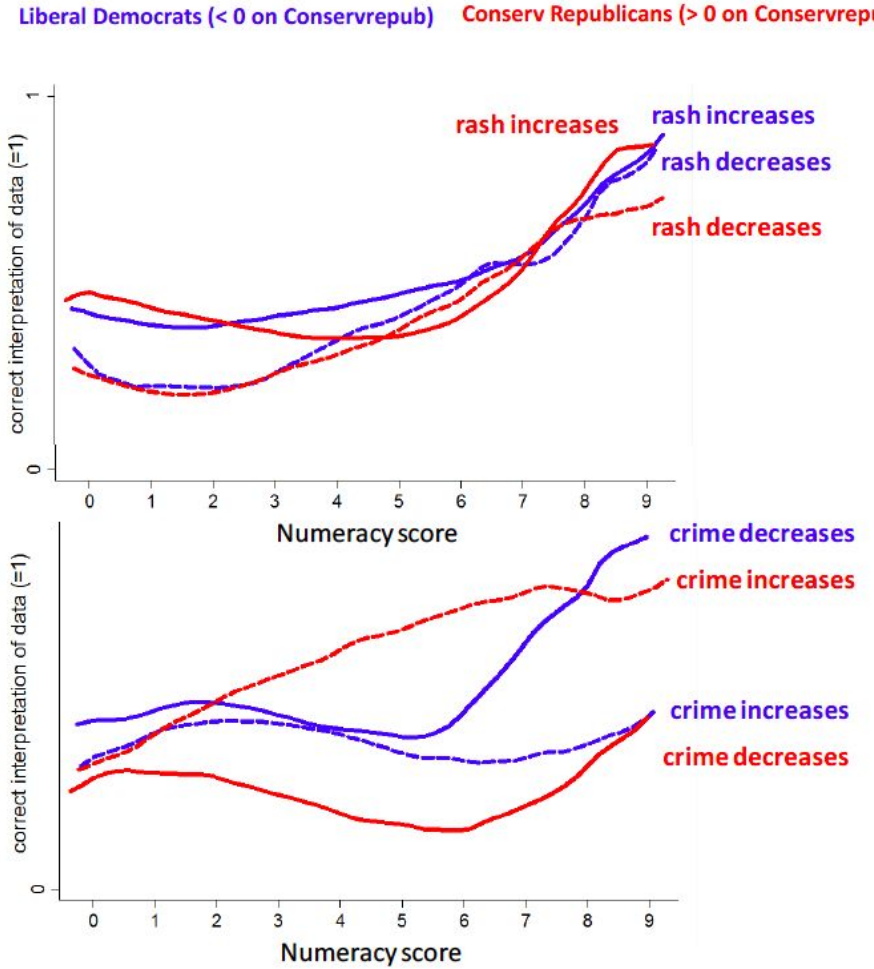


Figure 3. Experimental conditions. Subjects were assigned to one of four conditions. The conditions are identified by labels *(A)-(D)* in a manner that indicates the result or outcome of the experiment that is most supported by the data contained in the relevant table. The correct interpretation of the data was manipulated by varying the result specified by the headings above the columns.

What result does the study support?

- People who used the skin cream were more likely to get better than those who didn't.
- People who used the skin cream were more likely to get worse than those who didn't.

Kahan et. al. - Motivated Numeracy and Enlightened Self-Government



John was driving over the speed limit (about 40 mph in a 30-mph zone) in order to get home in time to . . .

Socially Desirable Motive
. . . hide an anniversary present for his parents that he had left out in the open before they could see it.

Socially Undesirable Motive
. . . hide a vial of cocaine he had left out in the open before his parents could see it.

Other Cause
Oil spill. As John came to an intersection, he applied his brakes, but was unable to stop as quickly as usual because of some oil that had spilled on the road. As a result, John hit a car that was coming from the other direction.
Tree branch. As John came to an intersection, he failed to see a stop sign that was covered by a large tree branch. As a result, John hit a car that was coming from the other direction.
Other car. As John came to an intersection, he applied his brakes, but was unable to avoid a car that ran through a stop sign without making any attempt to slow down. As a result, John hit the car that was coming from the other direction.

Consequence of Accident
John hit the driver on the driver's side, causing him multiple lacerations, a broken collar bone, and a fractured arm. John was uninjured in the accident.

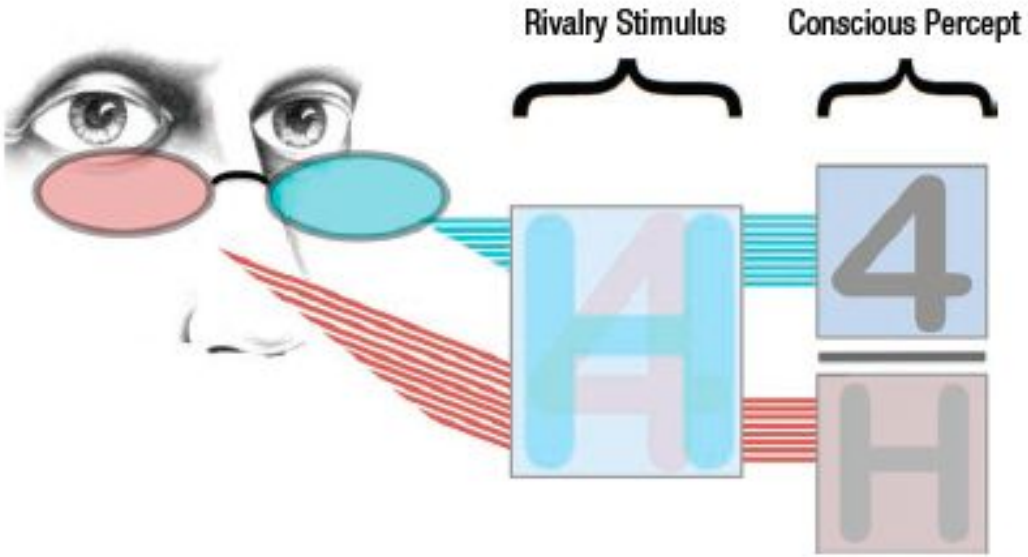
Complete the following sentence: The primary cause of this accident was ———.

To what extent was John the cause of the accident?

more the cause when the motive was socially undesirable

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

Dunning and Balcetis - Wishful Seeing: How Preferences Shape Visual Perception (2013)



Wishful Seeing

(Balcetis, Dunning, & Granot, 2012). We had participants wear colored goggles and presented the rivalry images for 300 milliseconds (see Fig. 2). The goggles filtered out different elements of the stimulus for each eye, such that one eye experienced a letter and the other a number. As we expected, participants were more likely to consciously see the image associated with financial gain than the image associated with loss, a pattern suggesting that desires influenced their preconscious categorization of ambiguous visual information to present to conscious awareness.

VALUE AND NEED AS ORGANIZING FACTORS IN PERCEPTION

BY JEROME S. BRUNER AND CECILE C. GOODMAN *
Harvard University

1947

PERSONAL VALUES AS SELECTIVE FACTORS IN PERCEPTION *

BY LEO POSTMAN *Indiana University* JEROME S. BRUNER *Harvard University* ELLIOTT McGINNIES *University of Alabama*

1948

Motivation biases perception

Balcetis & Dunning, 2006

Balcetis, Dunning & Granot, 2012

Subjective value determines initial dominance in binocular rivalry

Emily Balcetis [a,*], David Dunning [b], Yael Granot [a]
[a] Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA
[b] Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853, USA

It is concluded that most of the substantive contributions of the experiments reviewed **cannot be demonstrated to be related to perceptual variables...**

The S[ubject], in responding to a perceptual situation, **tends to respond in terms of the**

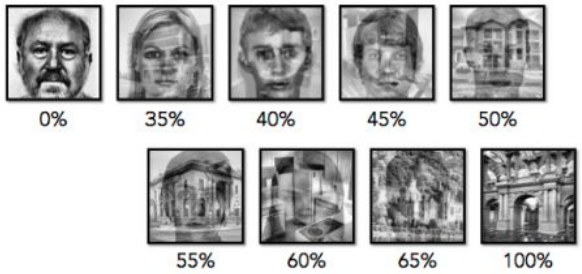
Motivation biases responses

experiments can readily lead to reports (of anything, including percepts) being contaminated by task demands, **wherein certain features of experiments lead subjects to adjust their responses** (either consciously or unconsciously) in accordance with their assumptions about the experiment's purpose.

Firestone & Scholl, 2016

Leong et. al. - Neurocomputational mechanisms underlying motivated seeing

Stimuli



Face/Scene Classification Task

Rewarded for correct classification

30 participants

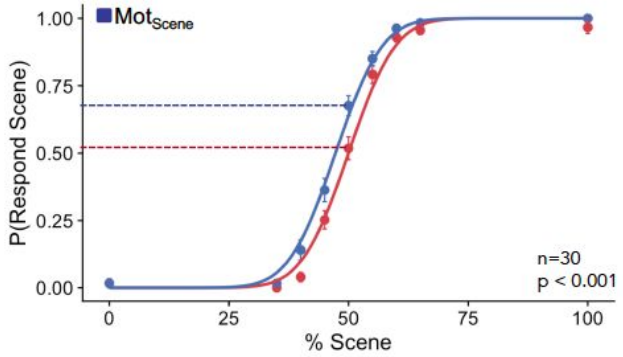
Motivation manipulation

Mot_{Scene}
- Win extra money if the image has more scene
- Lose some of their earnings if the image has more face

Mot_{Face}
- Win extra money if the image has more face
- Lose some of their earnings if the image has more scene

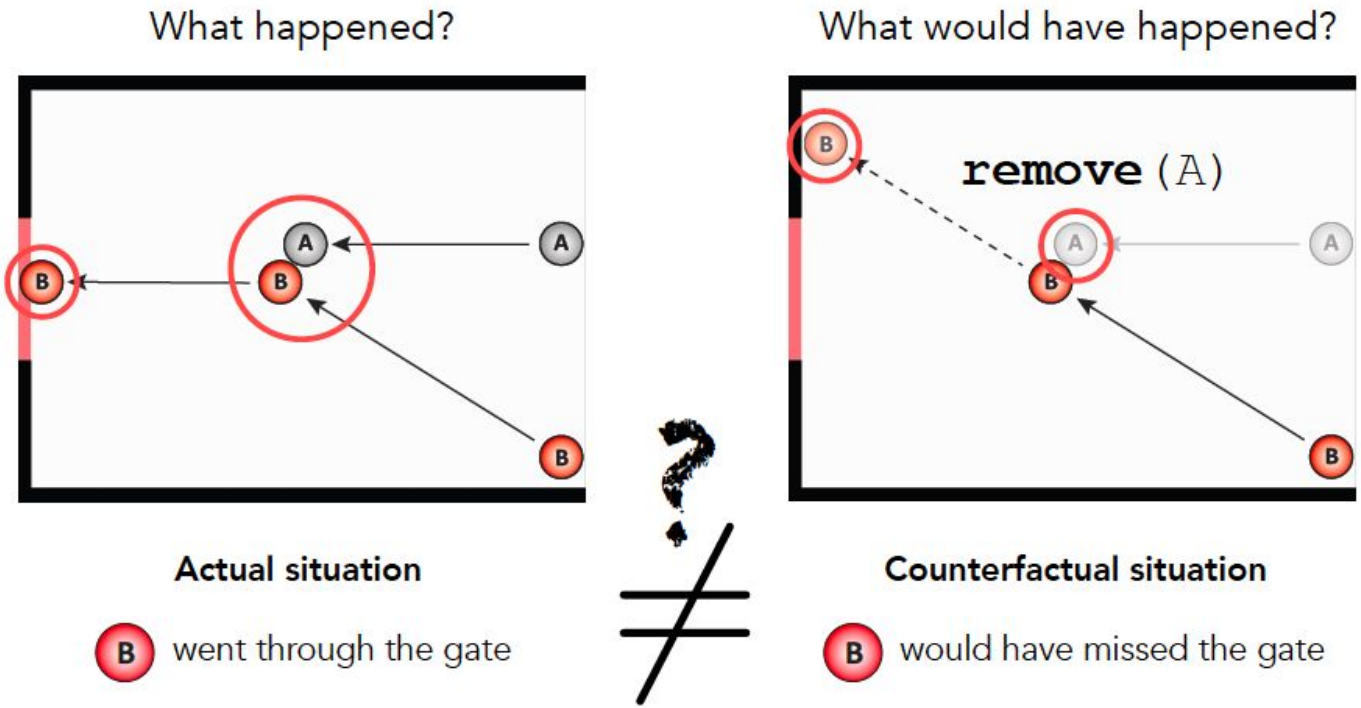
* Reward maximizing strategy was to ignore the motivation manipulation

Does motivation bias perceptual judgments?



Modeling results suggest that motivation biases both perceptual processing and responses.

Counterfactual Simulation Model

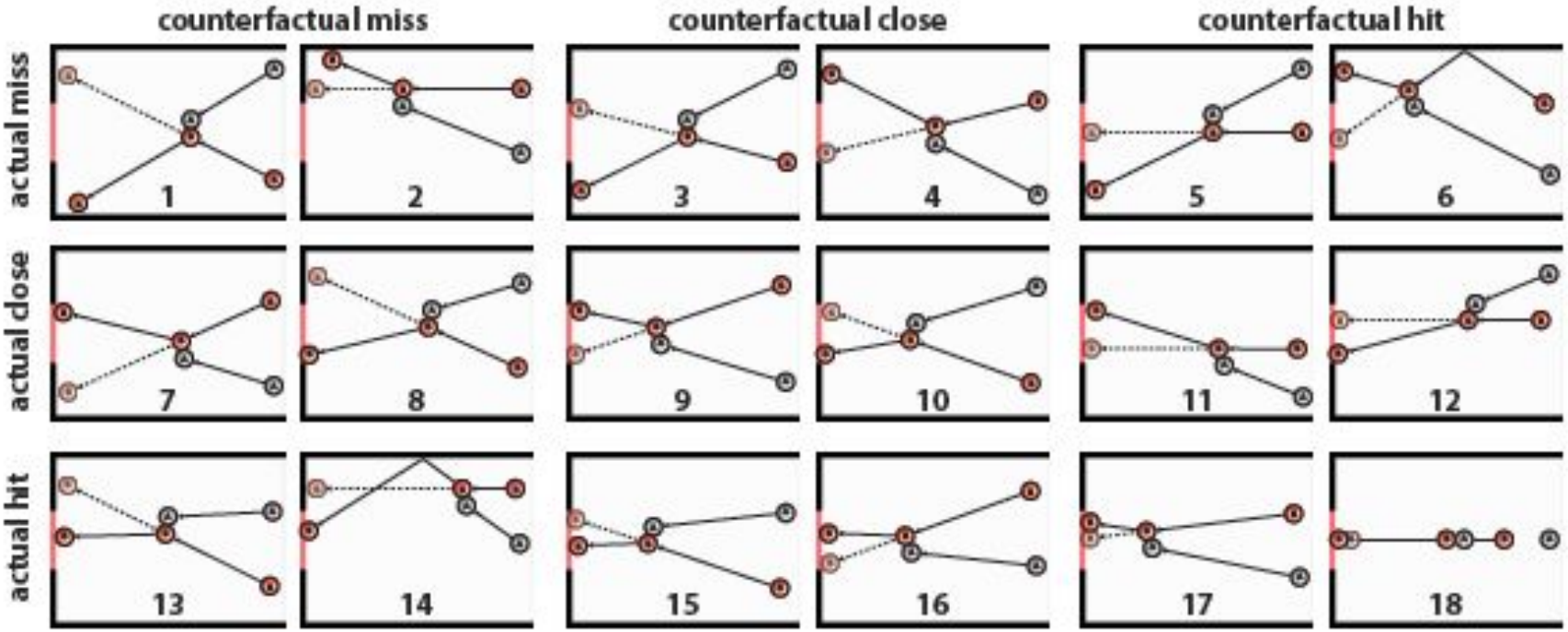


Gerstenberg, Goodman, Lagnado, & Tenenbaum (2012) Noisy Newtons: Unifying process and dependency accounts of causal attribution. *Cognitive Science Proceedings*

Gerstenberg, Goodman, Lagnado, & Tenenbaum (2014) From counterfactual simulation to causal judgment. *Cognitive Science Proceedings*

Gerstenberg, Goodman, Lagnado, & Tenenbaum (2015) How, whether, why: Causal judgments as counterfactual contrasts. *Cognitive Science Proceedings*

Evidence for CSM: Previous Experiment



Evidence for CSM: Previous Experiment

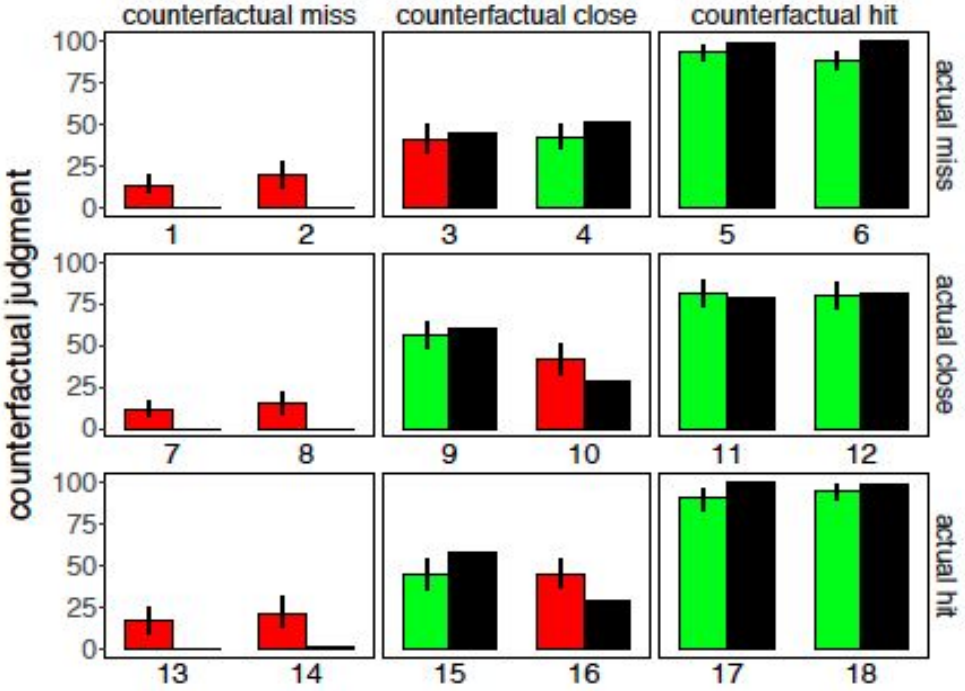
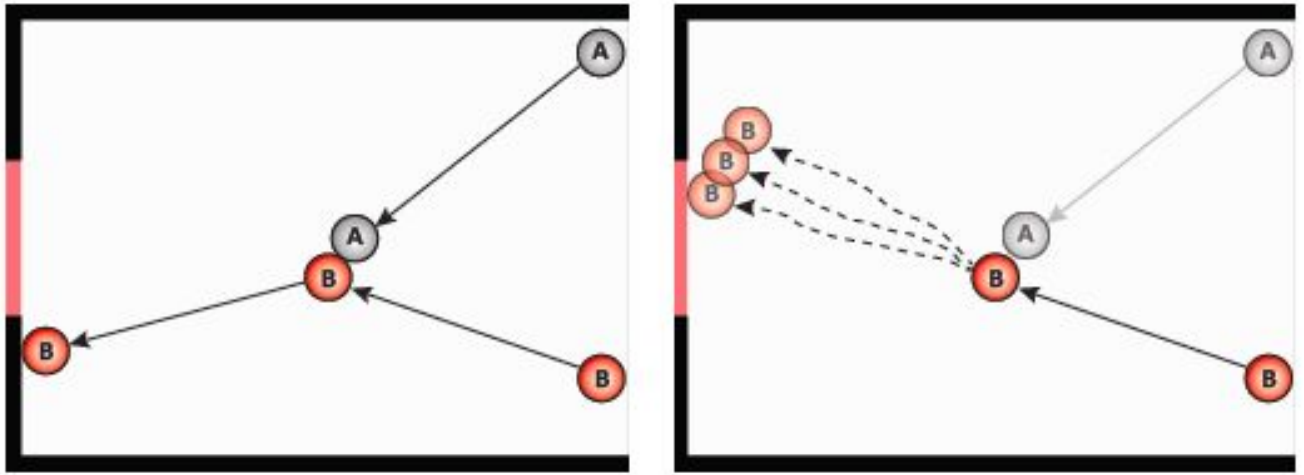


Figure 9. Mean counterfactual judgments (colored bars) together with the model predictions (black bars) of the best-fitting approximate simulation model. *Note*: Red bars indicate cases in which ball B would have missed. Green bars indicate cases in which B would have gone through the gate. 0 = "ball B would have definitely missed", 100 = "ball B would have definitely gone through the gate". Error bars indicate bootstrapped 95% confidence intervals.

Model captures human uncertainty better by introducing noise to ball B's counterfactual path



(a) actual situation

(b) three counterfactual simulations with small degree of noise

Evidence for CSM

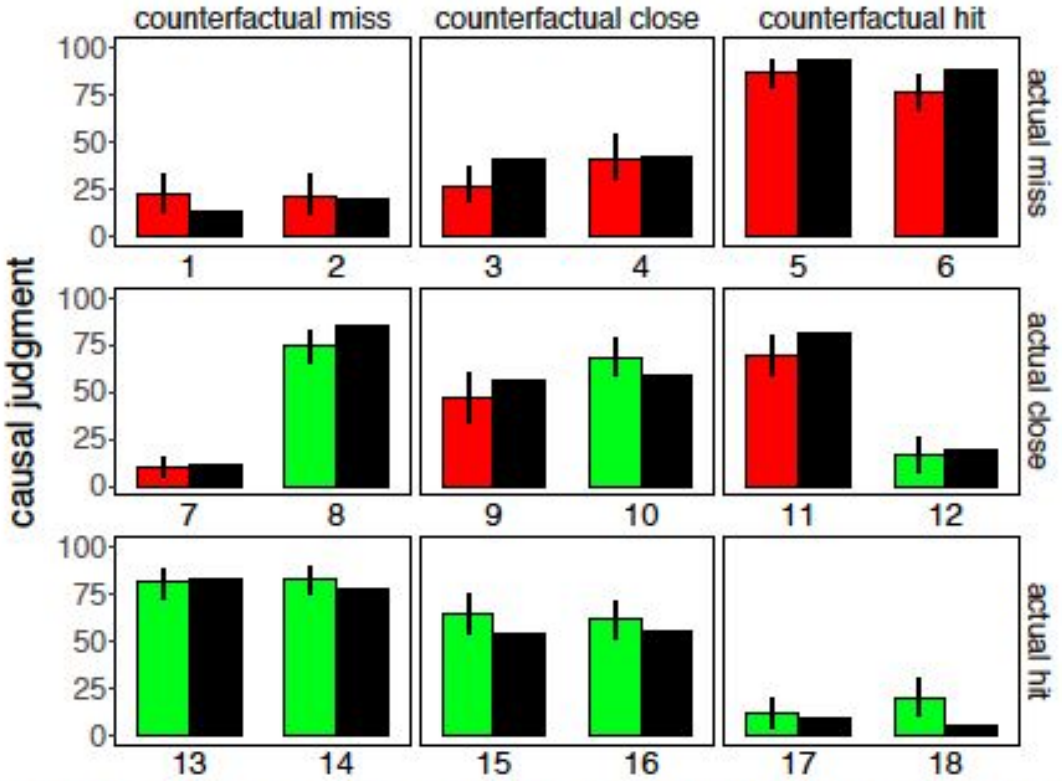
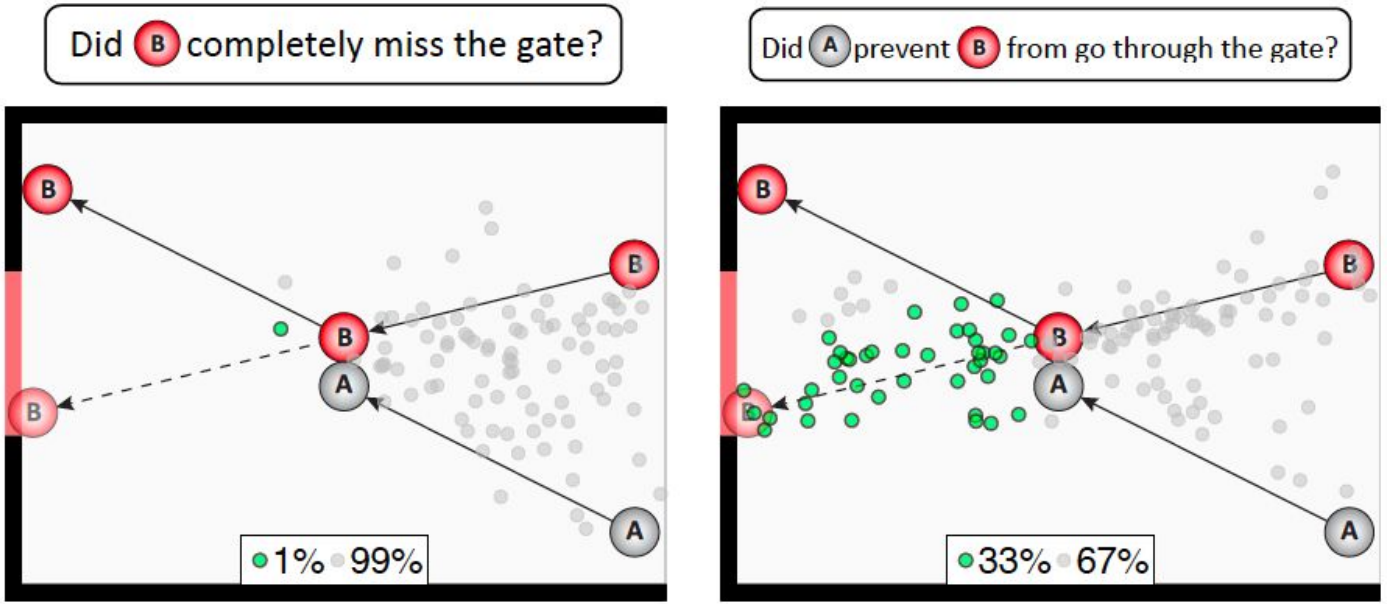


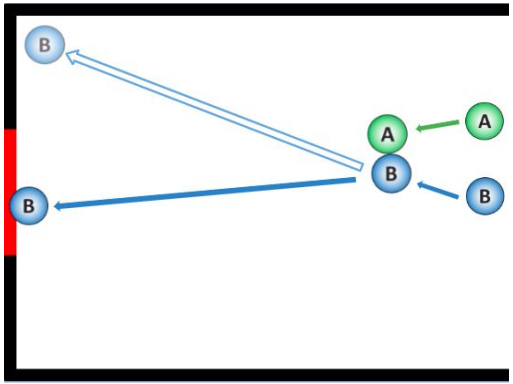
Figure 10. Causal judgments with model predictions. *Note*: Red bars are prevention judgments and green bars are causation judgments. Error bars indicate bootstrapped 95% confidence intervals.

Evidence for CSM: Spontaneous Counterfactual Simulation

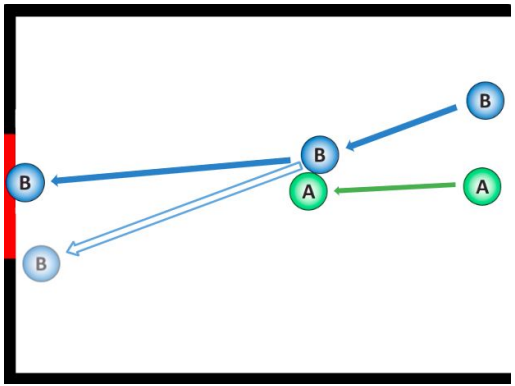


Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum (2017) Eye-tracking causality. *Psychological Science*

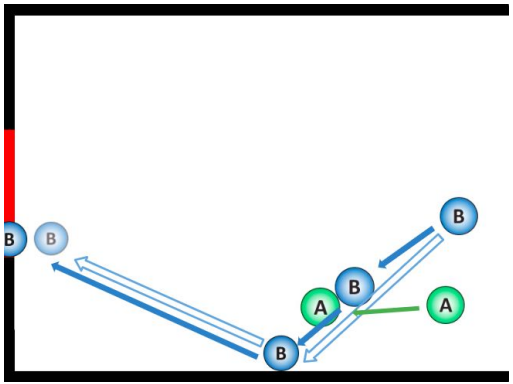
4 types of clips



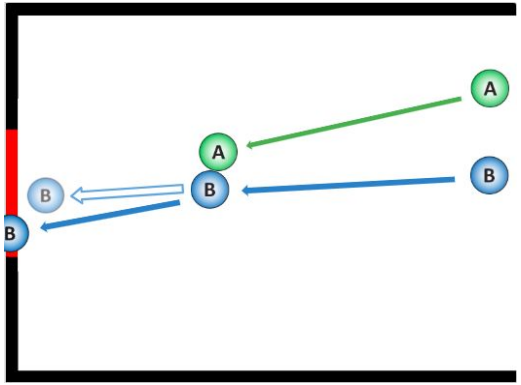
counterfactual: clear miss, causation: yes



counterfactual: close miss, causation: yes



counterfactual: close hit, causation: no

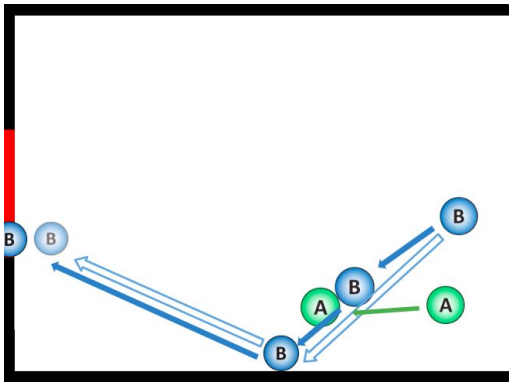


counterfactual: clear hit, causation: no

Probabilities of counterfactual hit: clear miss = [0,0.25), close miss = [0.25, 0.5), close hit = [0.5,0.75], clear hit = (0.75, 1]

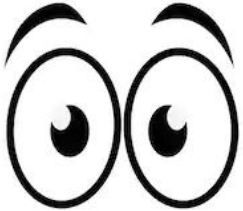
Experimental Approach

Did ball A cause ball B to go through the red gate?



Combining ambiguous perceptual stimuli and causal judgments

Responses (mturk)



Eye-tracking (in-lab)

Experiment 1: mturk (N = 46)

Did ball A cause ball B to go through the red gate?

NO

press f

YES

press j

Participants are rewarded 3 cents for each correct answer

Motivation Manipulation: between participants

Motivated for yes causation

- Win an extra 7 cents if ball A did cause ball B to go through the red gate
- Lose 7 cents if ball A did not cause ball B to go through the red gate

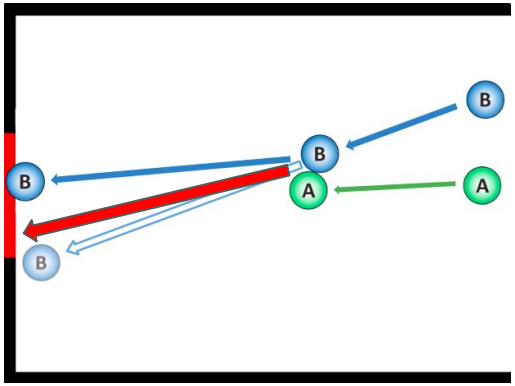
Motivated for no causation

- Win an extra 7 cents if ball A did not cause ball B to go through the red gate
- Lose 7 cents if ball A did cause ball B to go through the red gate

*Reward maximizing strategy ignores motivation manipulation and optimizes accuracy for +3 cents

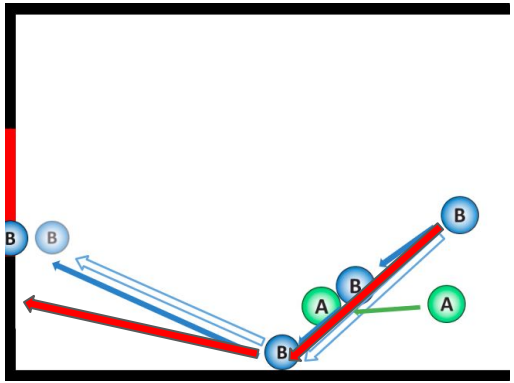
Hypotheses

Condition 1: motivated for no cause



ground truth: yes cause

Condition 2: motivated for yes cause



ground truth: no cause

Behavioral Predictions

more likely to say that A did not cause ball B to go through the gate

more likely to say that A caused ball B to go through the gate

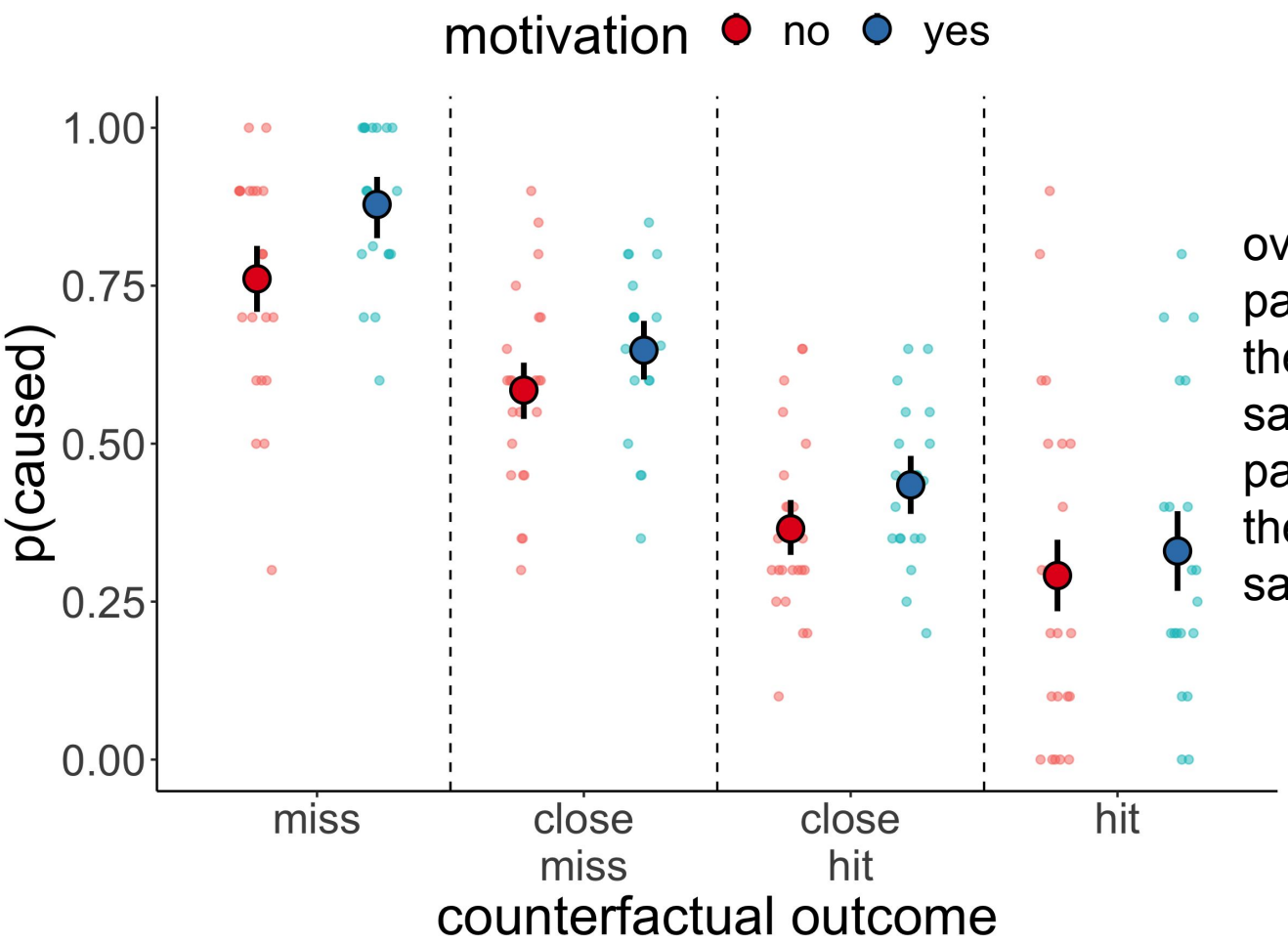
Eye-movement Predictions

more likely to simulate that B would have gone through the gate

more likely to simulate that B would have missed the gate

Results

Experiment 1: mturk



overall: participants in the yes condition say yes 56% and participants in the no condition say yes 49%.

Experiment 2: eye-tracking (in progress)

Did ball A cause ball B to go through the red gate?

Participants are rewarded **5 cents** for each correct answer

Motivation Manipulation: within participants

Motivated for yes causation

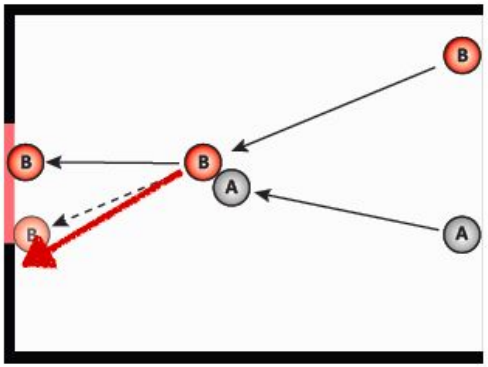
- Win an extra **20 cents** if ball A did cause ball B to go through the red gate
- Lose **20 cents** if ball A did not cause ball B to go through the red gate

Motivated for no causation

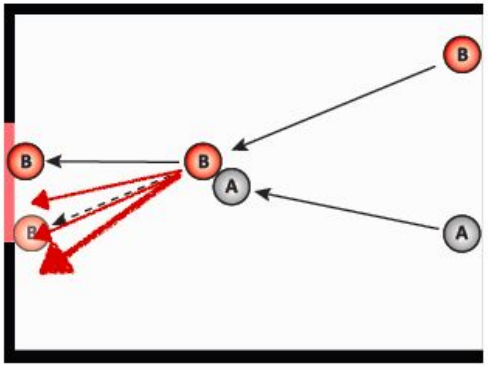
- Win an extra **20 cents** if ball A did not cause ball B to go through the red gate
- Lose **20 cents** if ball A did cause ball B to go through the red gate

*Reward maximizing strategy ignores motivation manipulation and optimizes accuracy for +5 cents

What aspects of causal judgment might motivations influence?



bias counterfactual simulations



weighting of simulations
number of simulations ...

we'll find out :)

Future Directions:

- ➔ Run ~30 eye-tracking participants
- ➔ Analyze eye-tracking data
  - ◆ Create video clips for helpful visualization
  - ◆ Create a plan and script to interpret eye saccades
  - ◆ See what we find
- ➔ Run mturk asking participants for only counterfactual judgments


Thanks!