The Language of Causation

Ari Beller*, Erin Bennett*, Tobias Gerstenberg
Stanford University

Introduction

The words we use matter.

Dr. Jones killed the patient.

vs

Dr. Jones caused the patient's death.

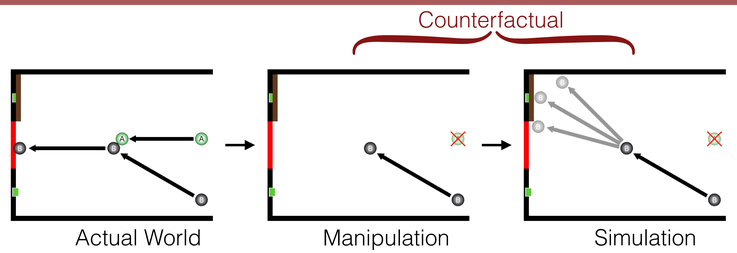
vs

Dr. Jones enabled the patient's death.

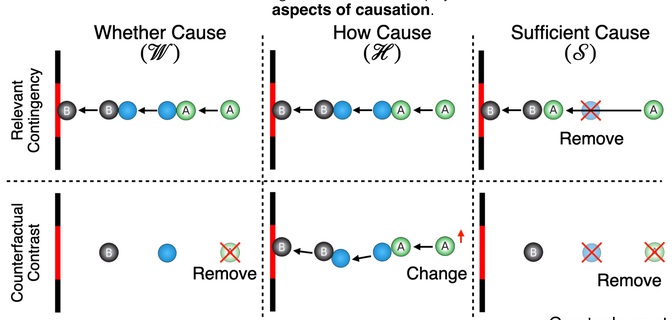
Rodriguez-Arias et al. (2020)

Model

1) Causal Inference



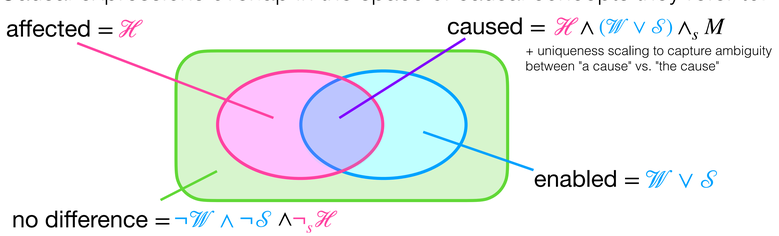
Different counterfactual tests in a generative model of physics allow us to asses different **aspects of causation**.



Gerstenberg et al. (2021)

2) Semantics

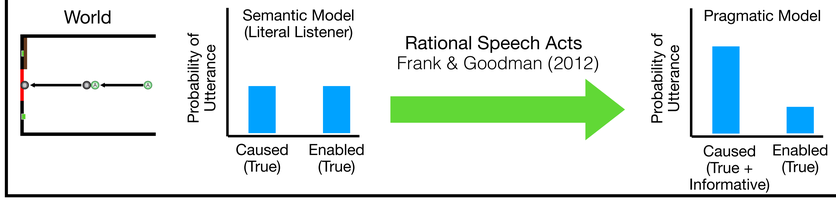
Causal expressions overlap in the space of causal concepts they refer to.



Model (cont.)

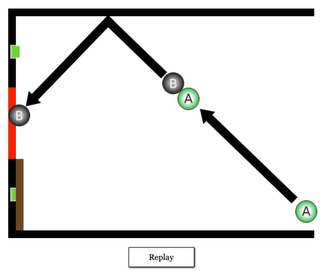
3) Pragmatics

For some worlds, multiple utterances are true. Pragmatics favors informative utterances.



Experiments

Experiment 1
Speaker Task



Which of the following sentences best describes the clip?

- Ball A **enabled** Ball B to go through the gate.
- Ball A **caused** Ball B to go through the gate.
- Ball A **affected** Ball B's going through the gate.
- Ball A **made no difference** to Ball B's going through the gate.

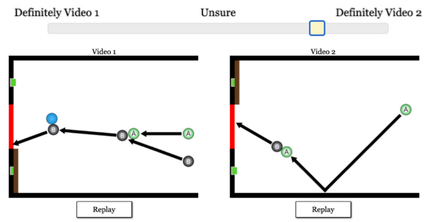
- 64 participants recruited on Amazon Mechanical Turk
- 30 video clips
- Within subject design

Experiment 2
Listener Task

The describer made the following selection:

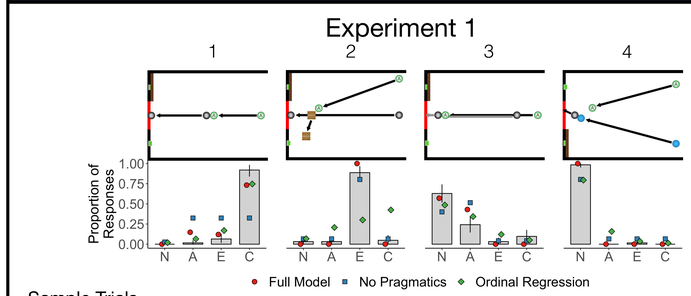
- Ball A **enabled** Ball B to go through the gate.
- Ball A **affected** Ball B's going through the gate.
- Ball A **caused** Ball B to go through the gate.
- Ball A **made no difference** to Ball B's going through the gate.

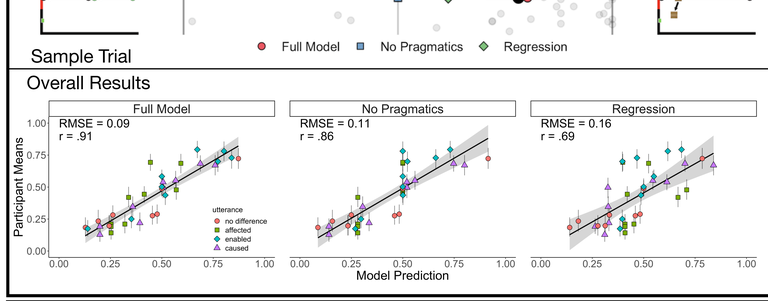
Which video do you think the describer saw?



- 50 participants recruited on Amazon Mechanical Turk
- 36 video clip pairings
- Within subject design

Results





Conclusions

- We implemented a model to investigate the connection between people's causal concepts and their causal language.
- In experimental tests, our full model outperformed two lesioned alternatives. This suggests causal knowledge, semantics, and pragmatics all play a role in people's causal language use.
- In future work, we plan to go beyond forced choice settings and account for more naturalistic descriptions.

References

- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336 (6084), 998–998.
- Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgment for physical events. *Psychological Review.*
- Rodríguez-Arias, D, Rodríguez López, B, Monasterio-Astobiza, A, Hannikainen, IR. How do people use 'killing', 'letting die' and related bioethical concepts? Contrasting descriptive and normative hypotheses. *Bioethics.* 2020; 34: 509– 518. https://doi.org/10.1111/bioe.12707