What's in a cause? Counterfactual relevance and hierarchical event structure in causal judgment

Jonathan F. Kominsky & Jonathan Phillips, Harvard University
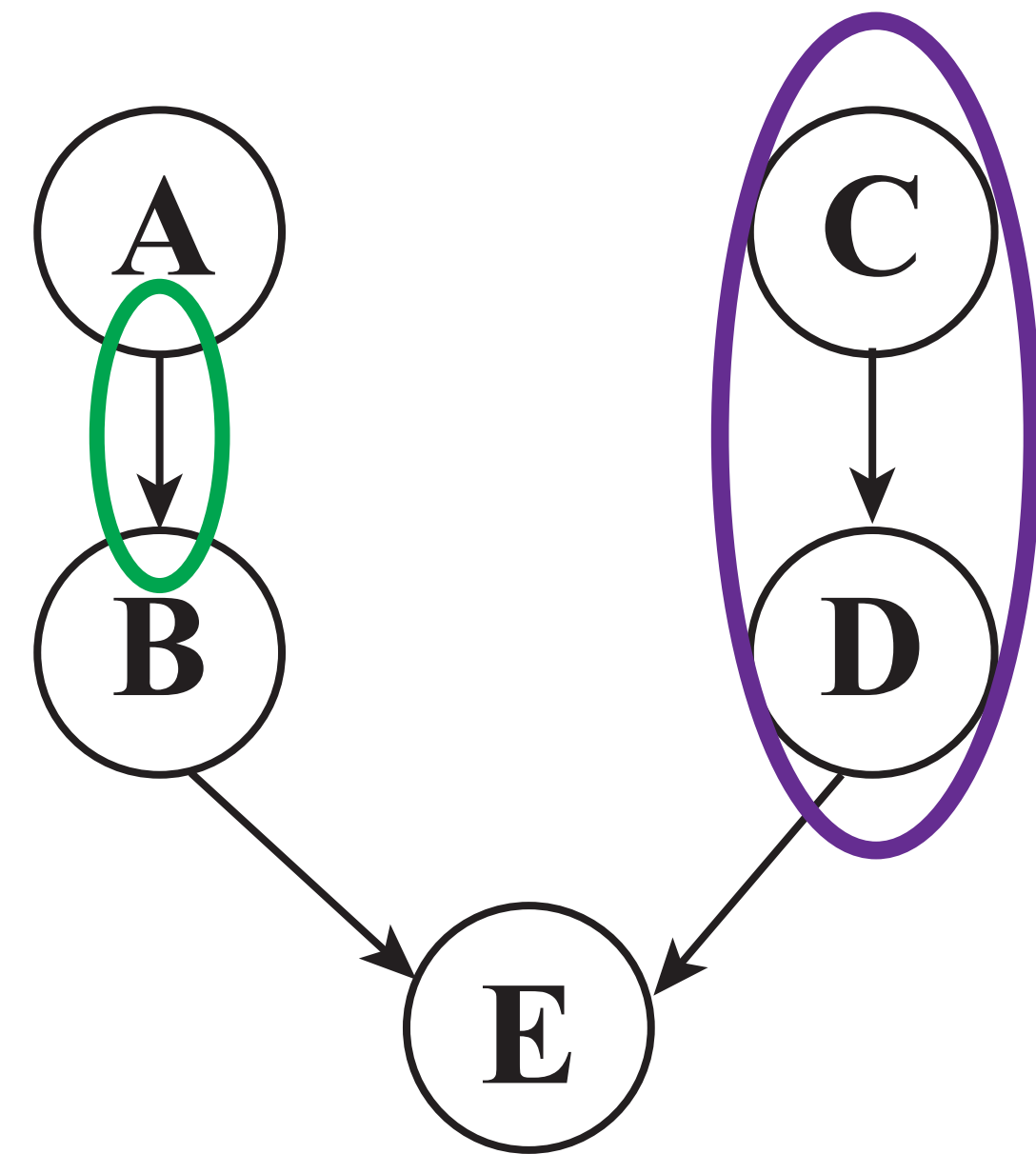
Introduction

Causal reasoning is often modeled as a graphical network (e.g., CBNs)

Well-studied:
How do we infer causal relationships between variables?
(Pearl, 2000; Griffiths & Tenenbaum, 2005; Sloman, 2005; many, many others)

Barely studied:
Which objects or events are represented as distinct causal variables in the first place?
(Halpern & Hitchcock, 2011)

Our question:

**How can we tell which events or objects are thought of as separate causal variables?**

Experiment 1

Inanimate objects and moral norms

Moral norms + mental state impact causal judgments of intentional **Agents**, but not the inanimate **Tools** they use. (Samland & Waldmann, 2016)

S&W (2016) results can be disambiguated by judgments of *counterfactual relevance.*
(Phillips et al., 2015)

N = 439
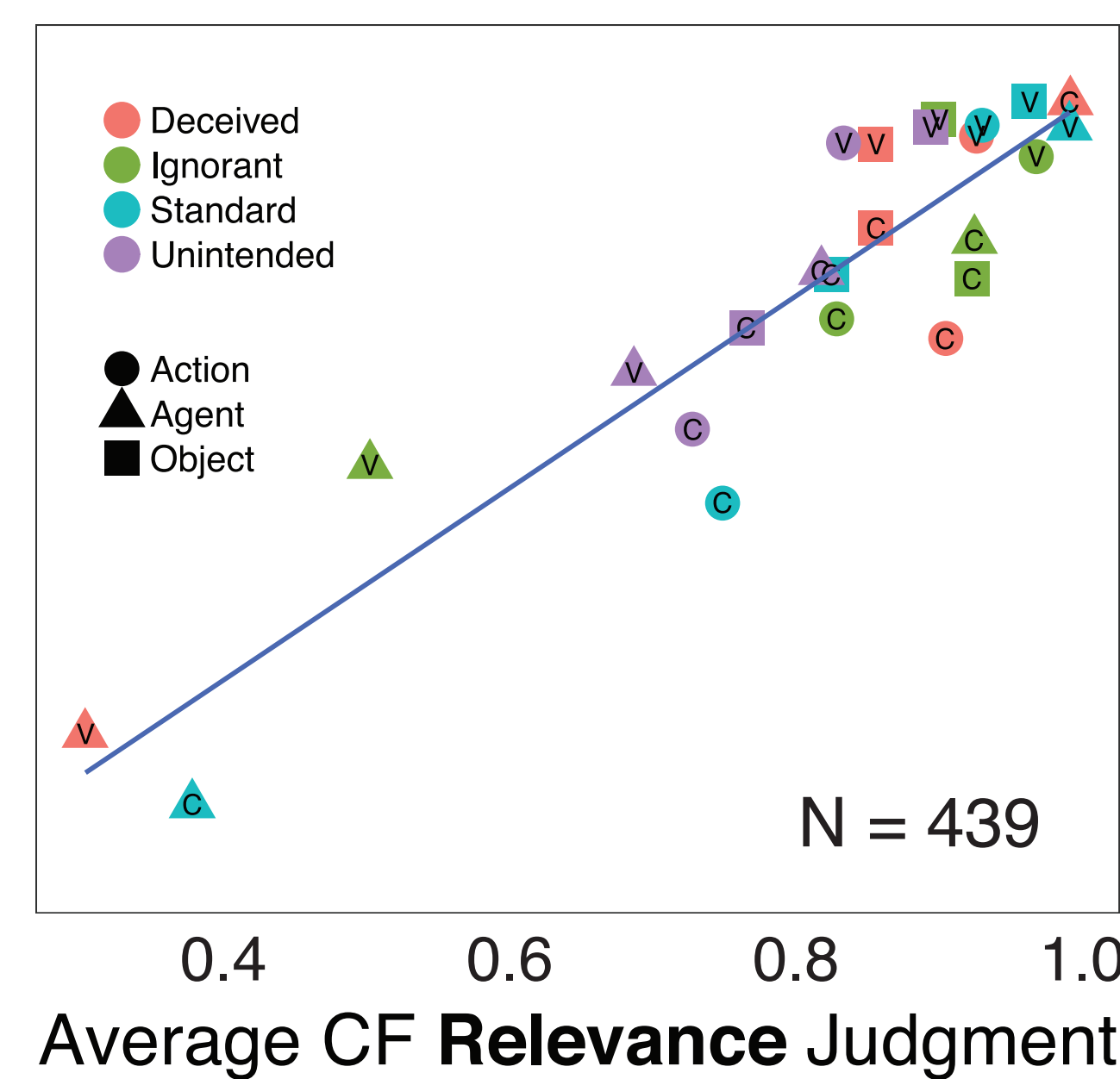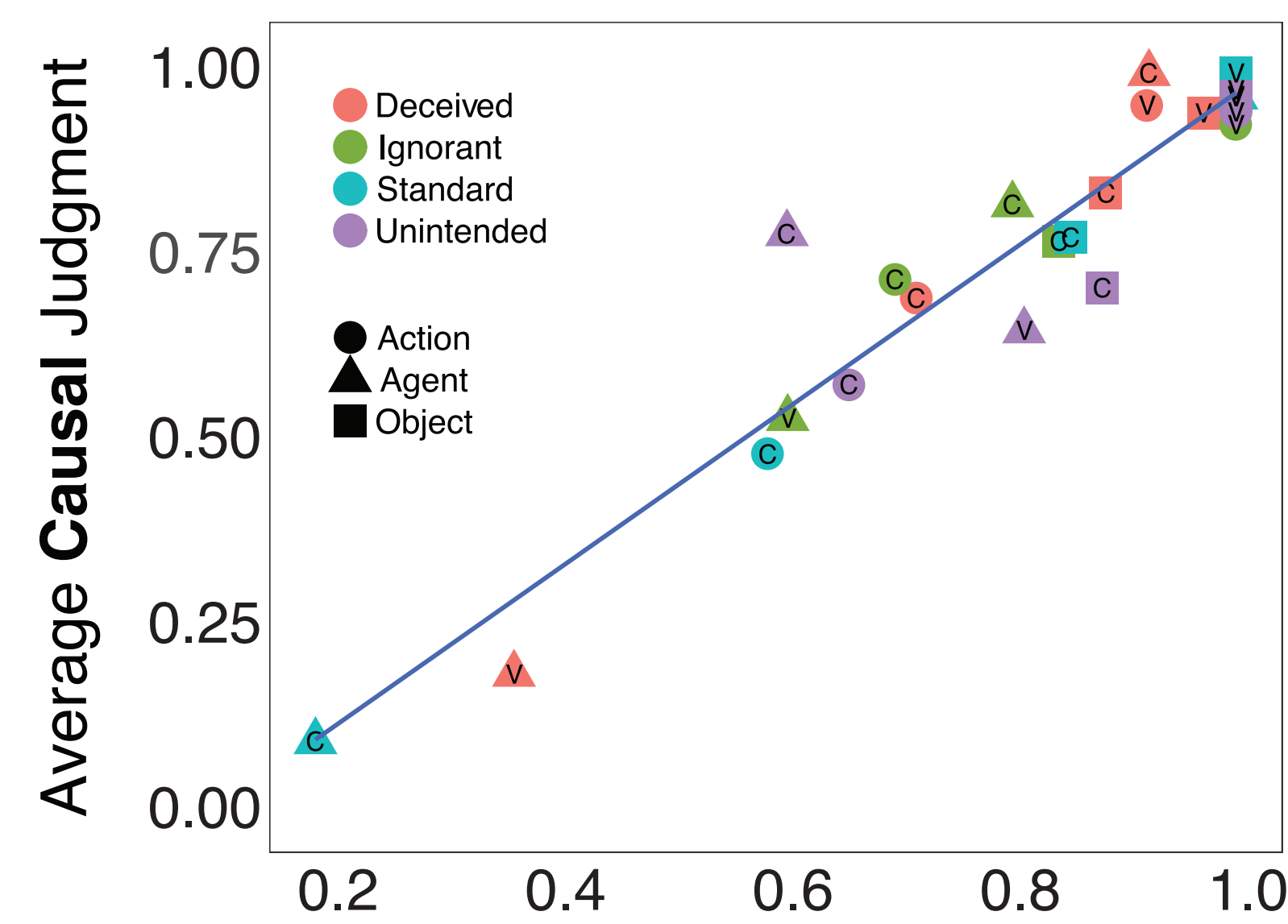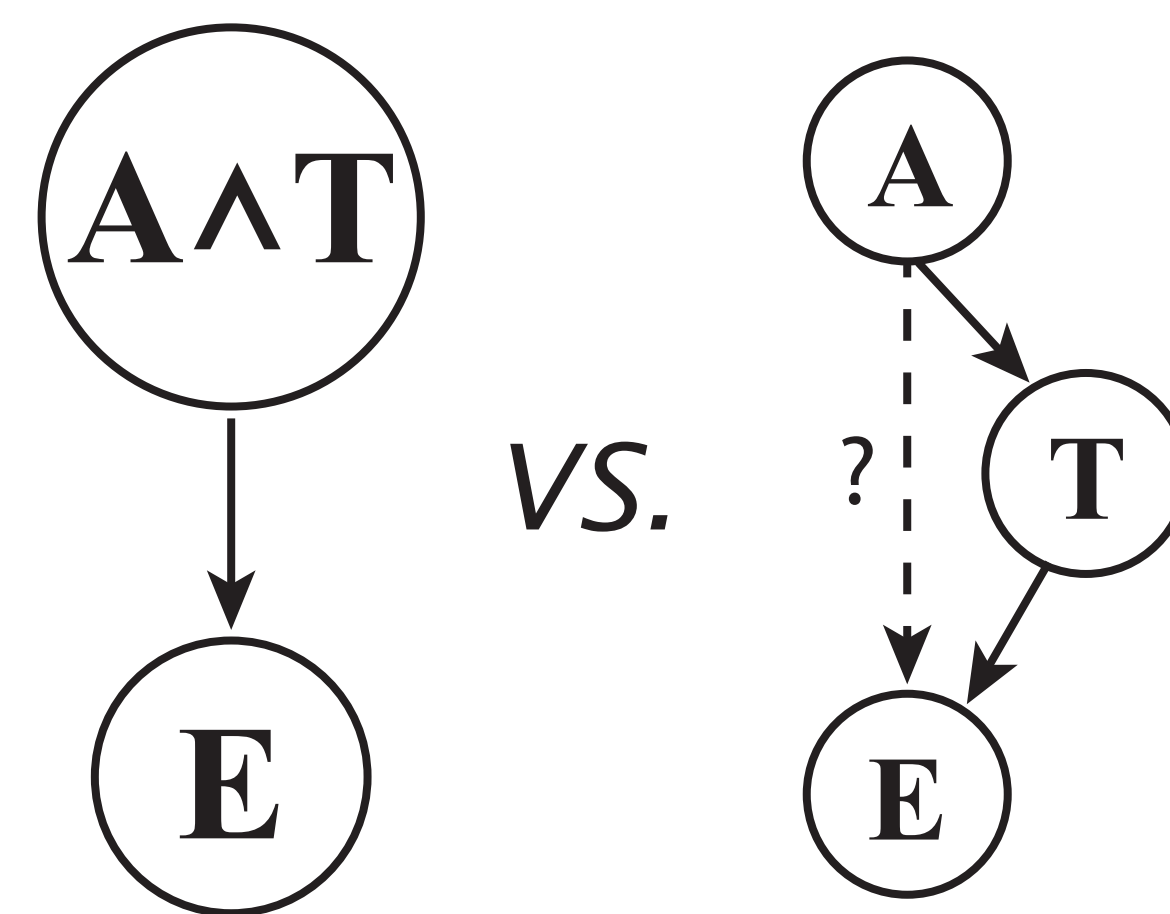
Experiment 2

Counterfactual induction influences causal judgments

"A department at the university recently got a new vending machine for dispensing office supplies. The machine has three levers, red, white, and black. Right now, pulling either the red lever or the white lever will produce a pencil. Pulling the black lever will produce an eraser.

Professor Smith and an administrator both need pencils. They both go to the machine. Professor Smith pulls the red lever, and the administrator pulls the white lever. Both get pencils.

Unfortunately, these were the last pencils in the machine. A student later needed a pencil for a test and went to the machine to get one, but there was a problem: there were no more pencils left in the machine."

Agent Counterfactual: "Consider what Prof. Smith could have done differently."

No Counterfactual: "Describe the story you just read"

Object Counterfactual: "Consider how the red lever could have functioned differently."
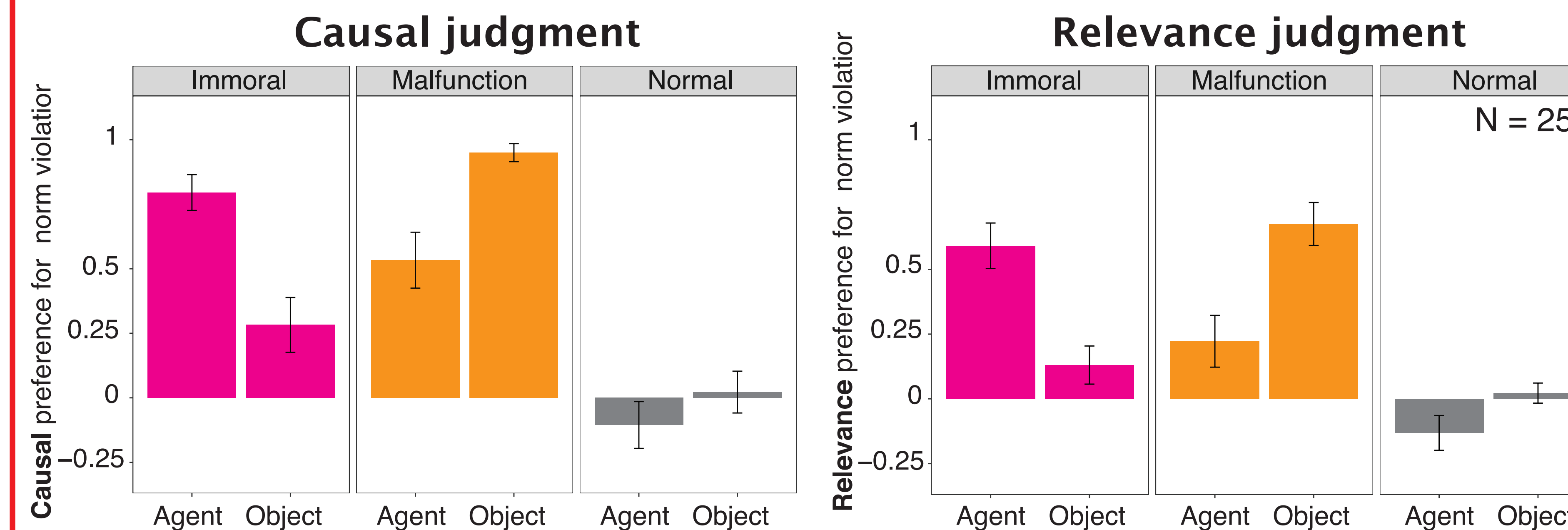
N = 423
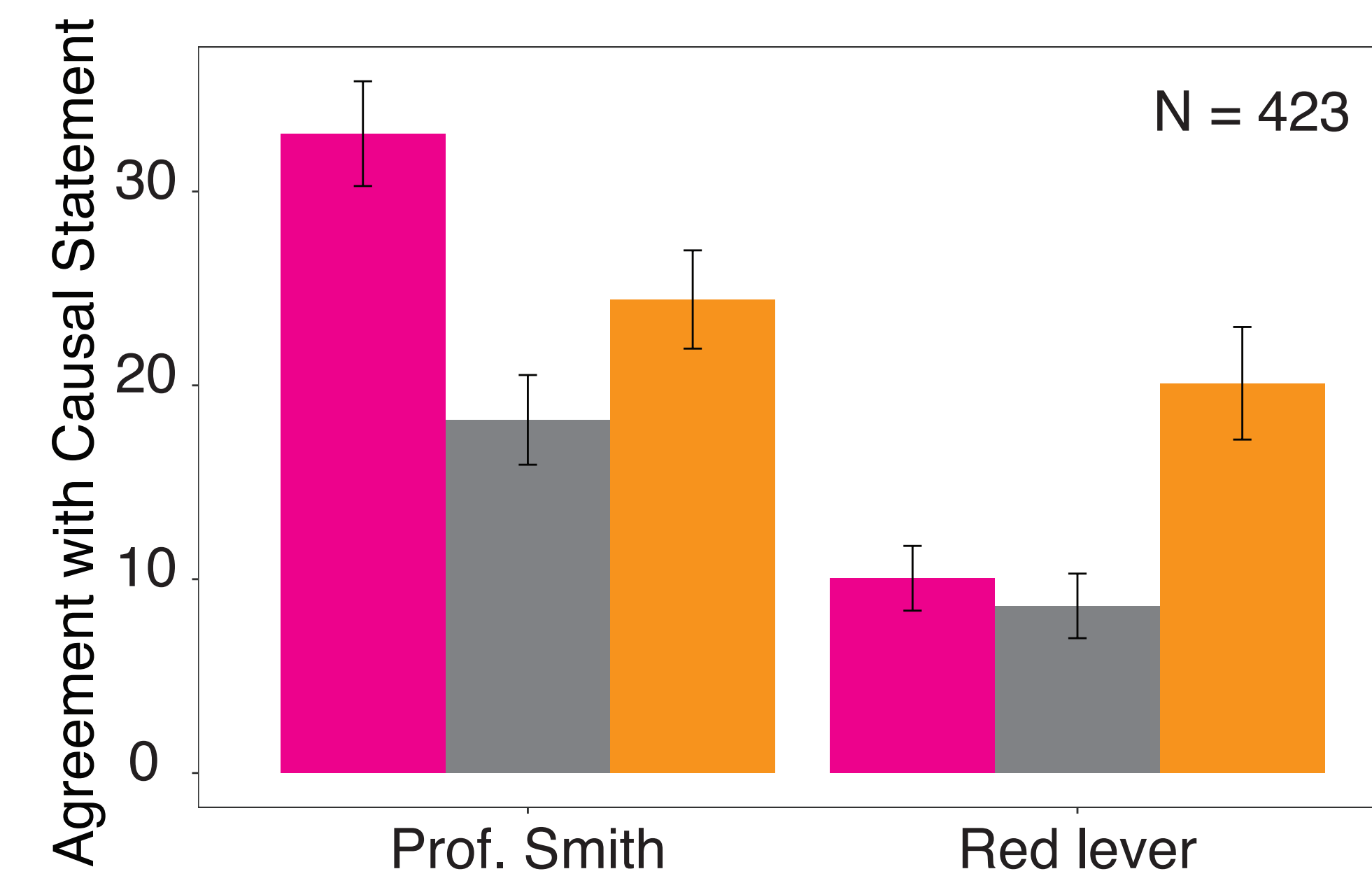
Experiment 3

Norms of *proper function*

Moral violation: "Professors are not allowed to get pencils from the vending machine, but administrative assistants are."

Malfunction: "The red lever produces erasers, but almost always malfunctions and also produces a broken pencil."
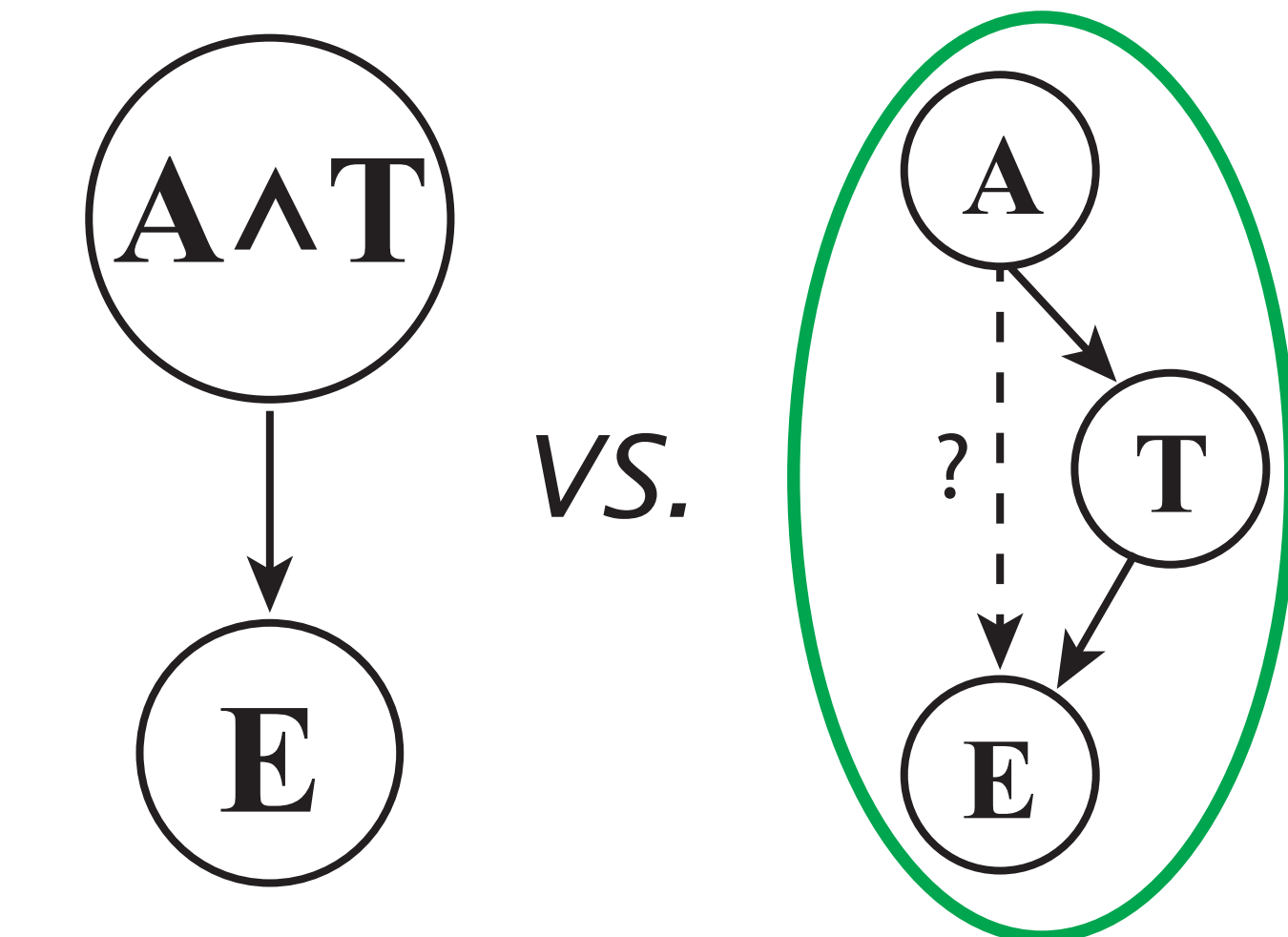
N = 258

Prescriptive norm violations can uncover dissociations between causal variables!

Conclusions

1. Norm violations influence causal judgments via *counterfactual relevance.*
(Icard et al., 2017; Kominsky et al., 2015; Phillips et al., 2015)

2.

3. A curious asymmetry: Factors affecting the relevance of agent counterfactuals do not impact judgments of objects, but factors affecting the relevance of objects affects judgment of agents!

3a. Further work has found this does not occur with hierarchical relationships of two agents (i.e. tool replaced with another agent), but does with two inanimate objects.
(Phillips & Kominsky, in prep)

References

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*(4), 334-84.

Halpern, J. Y., & Hitchcock, C. (2011). Actual causation and the art of modeling. ArXiv Preprint ArXiv:1106.2652.

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

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition, 137*, 196-209.

Pearl, J. (2000). Causality : Models, reasoning, and inference. Cambridge, U.K. ; New York: Cambridge University Press.

Phillips, J., Luguri, J. B., & Knobe. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities . *Cognition, 145*, 30-42.

Samland, J., & Waldmann, M. R. (2016). How prescriptive norms influence causal inferences. *Cognition, 156*, 164-176.

Sloman, S. A. (2005). Causal models: How we think about the world and its alternatives. New York: Oxford University Press.

Gear images created by Freepik