Variability in Causal Reasoning

Ivar Kolvoort[1], Zachary Davis[1], Leendert van Maanen, & Bob Rehder
1 equal contribution. IR: i.r.kolvoort@uva.nl, University of Amsterdam. ZD: zach.davis@stanford.edu, Stanford University

Background

Causal reasoning is a core cognitive phenomenon, it impacts our beliefs, attitudes, judgements and actions. However, **our current understanding of the cognitive processes that produce causal judgments is limited.**

How can we evaluate the processes that generate causal judgments? The predominant approach is to assess the predictions of multiple models against the average judgments of participants. This approach is principled and effective, but utilizing only **averaged data has not been able to convincingly identify the best model** out of the multitude that have been proposed (Rehder, 2014, 2018; R&H, 2016).

Distributions instead of means

- **To overcome the unidentifiability of theories of causal reasoning, we analyze full response distributions instead of just mean responses.**
- A few studies have remarked on the considerable variability in human causal judgments (D&R, 2020; Rehder, 2014; R&H, 2016). However, it is unclear to what extent that represents within- or between-participant variability.
- One major methodological hurdle in analyzing within-participant variability has been the elicitation of independent judgments for repeated causal queries.

Aims

As this is **the first study to analyze response distributions of causal judgments**, we aim to answer some foundational questions.

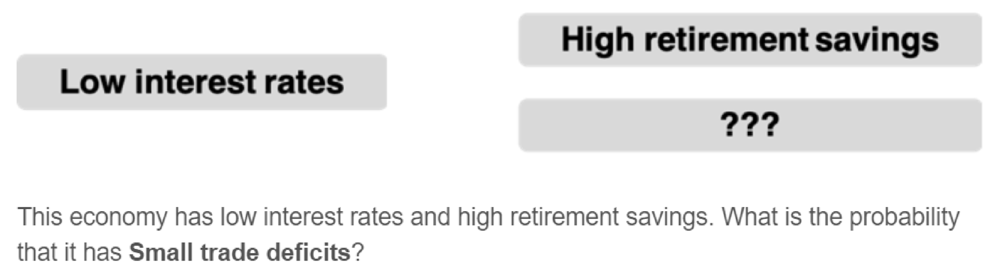
- Establish whether **there is meaningful within-participant variability**
- **Compare variability across inference types**:
  - Predictive vs diagnostic reasoning
  - Effect of conditional information
- Test if individual level variability is related to the systematic **violation of Markov independence**
- Provide a **comparison of the observed variability against qualitative predictions** of existing models

Experiment

- 37 participants (recruited via Prolific.co, 8 removed based on comprehension checks)
- Domains and descriptions of causal relationships from Rehder (2014)

Procedure

1. Subjects first studied several screens with a cover story and a description of each domain's binary causal variables and their causal relationships, which formed a **common cause network**
   - Base rates of all variables are 50%
   - Each cause produced its effect "75% of the time"
2. Inference test: each trial presented the values of 1 or 2 of the variables and asked participants to predict the state of another.



Design

We chose inference types to study effects of direction and conditional information, resulting in (2x3=)6 inference types.

Table 1: Inference Types and Normative Answers

	Diagnostic	Predictive
Consistent	$P(Y \mid X_i=1,X_j=1) = 94\%$	$P(X_i \mid Y=1,X_j=1) = 80\%$
Incomplete	$P(Y \mid X_i=1) = 80\%$	$P(X_i \mid Y=1) = 80\%$
Inconsistent	$P(Y \mid X_i=1,X_j=0) = 50\%$	$P(X_i \mid Y=1,X_j=0) = 80\%$

Repeated independent measurements:

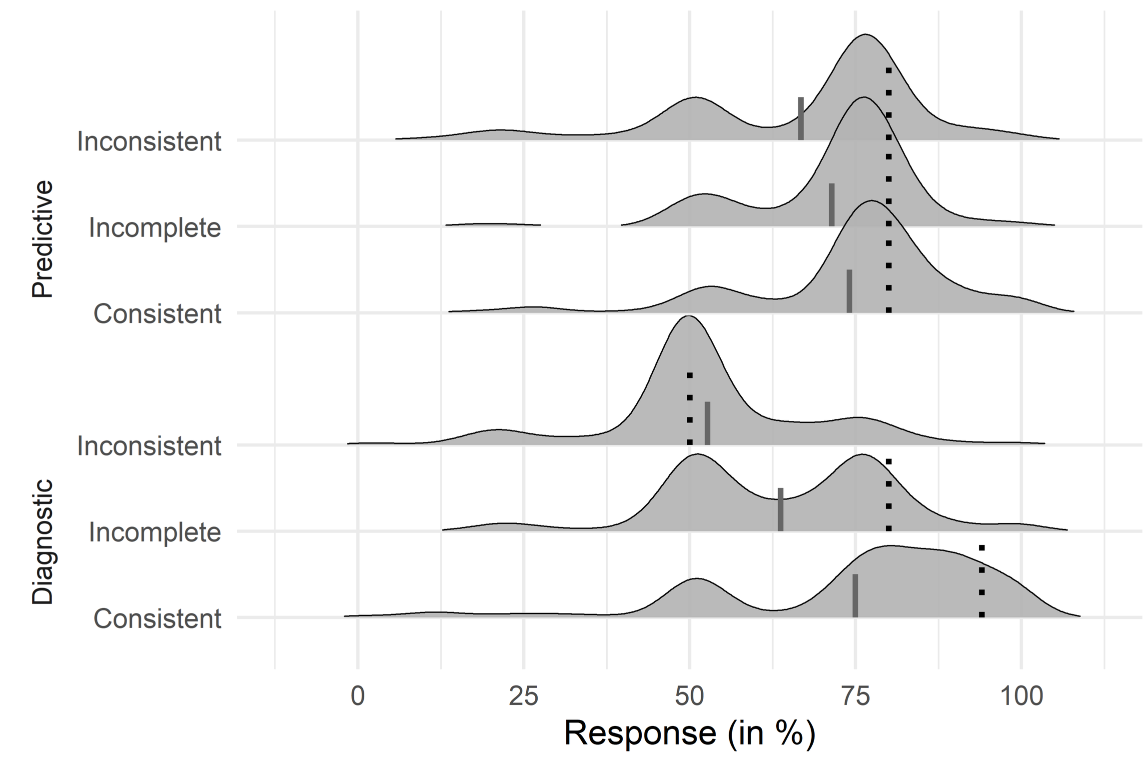
To get multiple measurements of a single inference type we collapsed over:

- 5 different domains
- Symmetry between the two effects
- Absence and presence P(Y=1) = 1- P(Y=0)

Resulting in **20 measurements of each inference** type, and 120 queries per participant.

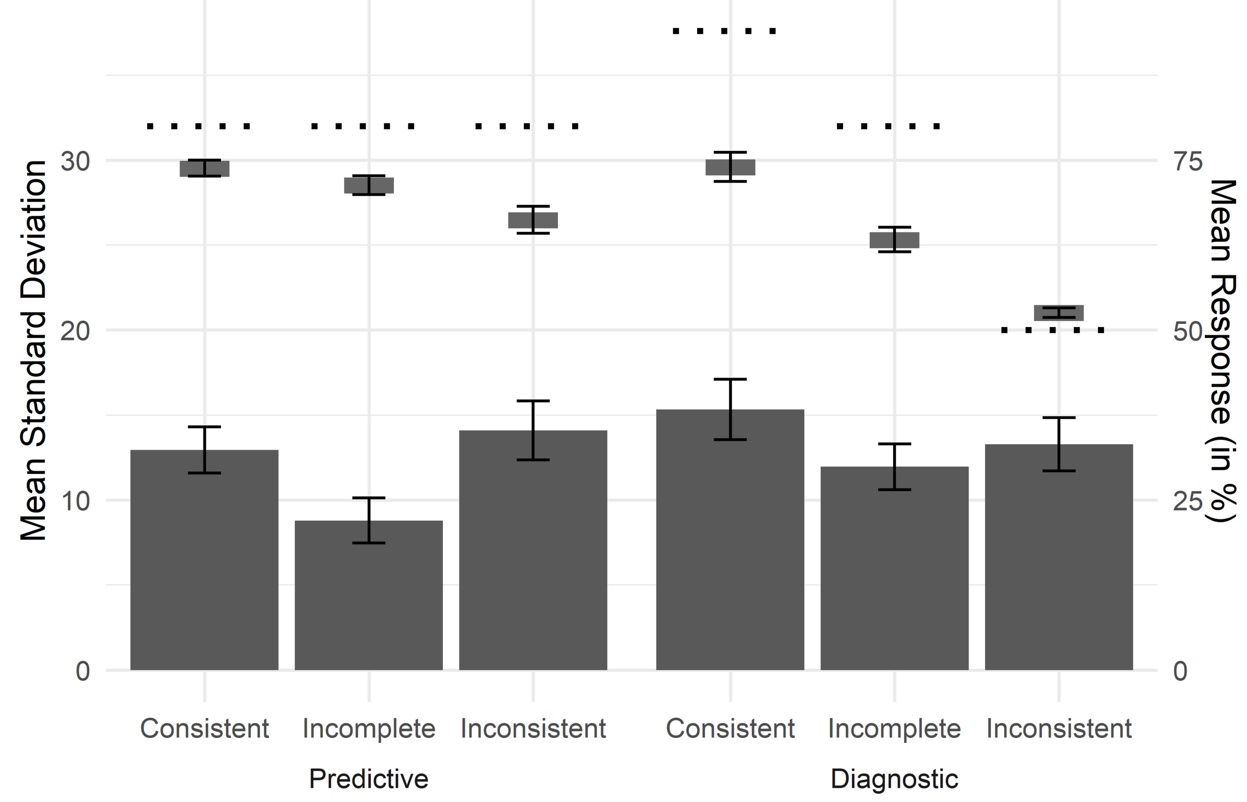
Results

Overall response distributions



Dotted lines are normative responses, solid lines indicate mean responses. **Individual response distributions are often multi-modal.** There are **spikes at 50%**, largest for inconsistent, and smallest for consistent information.

Variability per inference type

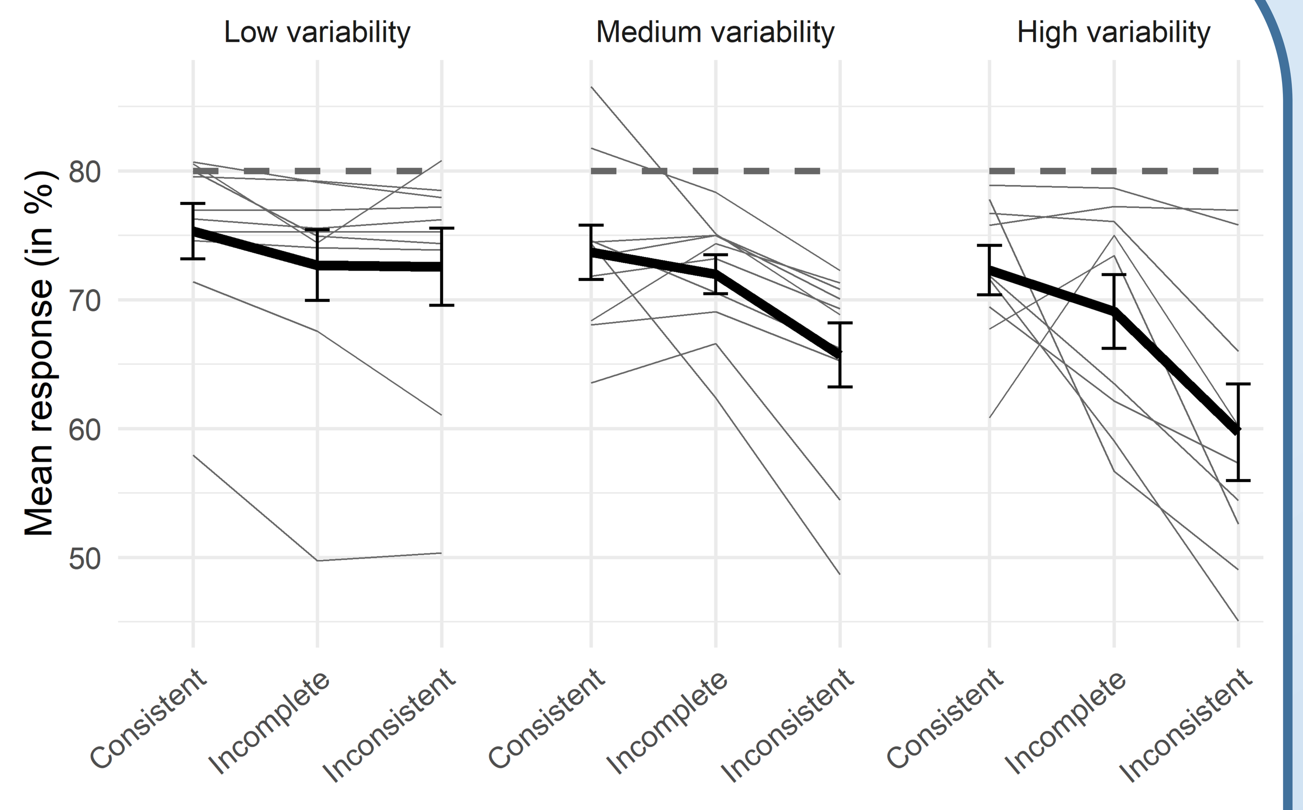


Within-participant variability (bars) does not follow mean response (floating boxes) and **varies systematically** per inference type:

- **Variability is lower for inferences with incomplete information.**
- **Variability is higher for diagnostic inferences.**

These findings indicate that the observed within-participant **variability reflects (at least partly) a decision-making process**, and not just noise.

Markov violations



More variable participants committed larger Markov violations. A common process might drive both Markov violations and a part of the observed variability

Sources of variability

- **Motor or general task noise cannot explain the variability**, as it varies systematically and is multi-modal.
- **Distributions are not centered on the normative response**, counter to predictions of e.g. the Beta inference model (R&H, 2016)
- **Uncertainty about the parameters of the causal network is unlikely as a source.** It could explain increased variability for diagnostic inferences, but cannot explain other findings.
- **Default responding might explain spikes at 50%.** A possible explanation of changes in spikes might be that guessing is more likely with more ambiguous information.
- **The Mutation Sampler** (D&R, 2020) can explain the changing spikes and predicts within-participant variability.

Conclusions

- Within-participant variability in causal reasoning can be probed experimentally and is related to the type of inference
- Variability can't be explained by simple additions to normative CGM model
- Computational models need to account for (systematic variation in) variability

References:
D&R: Davis & Rehder (2020) A Process Model of Causal Reasoning. Cog. Sci.
Rehder (2014) Independence and Dependence in Human Causal Reasoning. Cog. Psy.
Rehder (2018) Beyond Markov: Accounting for independence violations in causal reasoning. Cog. Psy.
R&H: Rottman & Hastie (2016) Do People Reason Rationally about Causally Related Events? Cog. Psy.