UNIVERSITY OF LÜBECK
INSTITUTE OF INFORMATION SYSTEMS

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CHAI
Humanities-Centered AI

# Predicting How People Update Beliefs

## Individual-, Subgroup-, and Population-Level Dynamic Bayesian Networks

Rebecca von Engelhardt[1] Malte Luttermann[2]

## 1. At Which Level Should We Model Belief Updating?

People revise beliefs about others from new evidence, but differ in *when* and *how*. Any model of belief updating must be fitted at some level of aggregation:

more generalisable to unseen people

| Population | Subgroup | Individual |
|---|---|---|
| one model for whole sample | one model per cluster | one model per person |

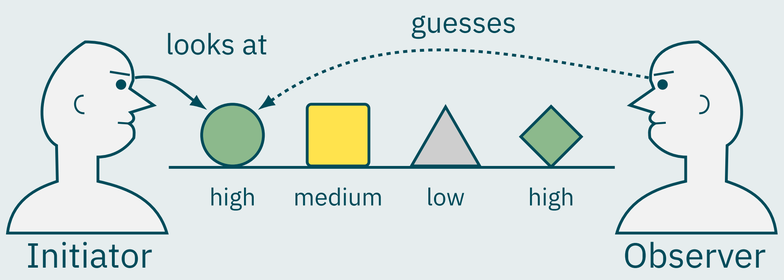
captures more individual differences

Population models *obscure* individual differences, individual models are *noisy* and hardly generalise to new people — clustering similar individuals is a principled compromise.

**Question**: Can cluster-based models predict belief updating as accurately as individual models while staying generalisable to unseen individuals?

## 2. Task: Cooperative Gaze-Following

An initiator looks at one of several objects and an observer has to guess which one. Correct guesses earn team points — more for high-value objects — which creates a cooperative context:

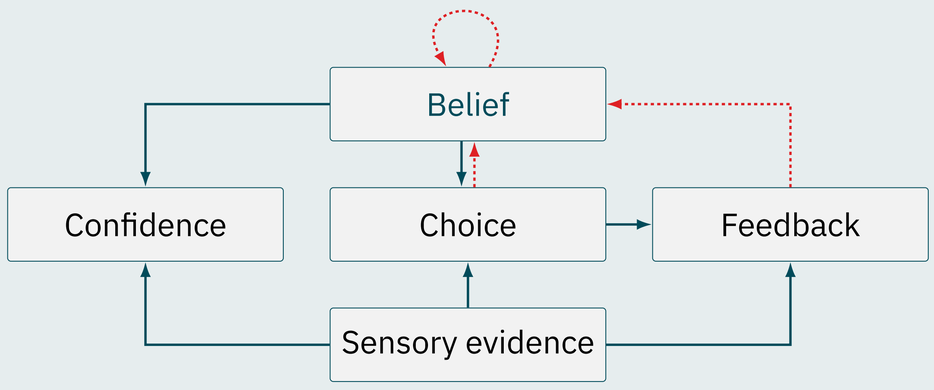



A confederate initiator manipulates cooperativeness across blocks:

- **Congruent**: mostly high-value objects (cooperative)
- **Incongruent**: more low-value objects
- **Random**: value chosen at random

## 3. Dynamic Bayesian Network (DBN) and Clustering

Each trial is one time slice with five variables linked by CPTs; solid edges act *within* a trial $t$, dashed edges carry to the *next* trial $t{+}1$.




All participants share this same known graph, so differences live only in the CPTs. We simulate $N = 100$ participants from four latent groups of updaters:

- **Choice-based**: update mainly after correct choices
- **Insecure**: similar, but noisier / more switching
- **Sceptical**: update slowly in positive directions
- **Persistent**: barely update at all

As the graph is fixed, clustering reduces to the CPT space: we fit one DBN per person and use the RMSE between their CPTs as distance for clustering.

## 4. Predicting Belief Updating in New Trials

We predict 180 new trials from the fitted CPTs at all three levels and average over 1000 forward-sampled runs (red = mean ± SD; black/grey = real data).




Individual (top) and cluster (middle) predictions track the real trajectories closely; the sample-level model (bottom) misses the subgroup-specific pattern.

## 5. Model Comparison

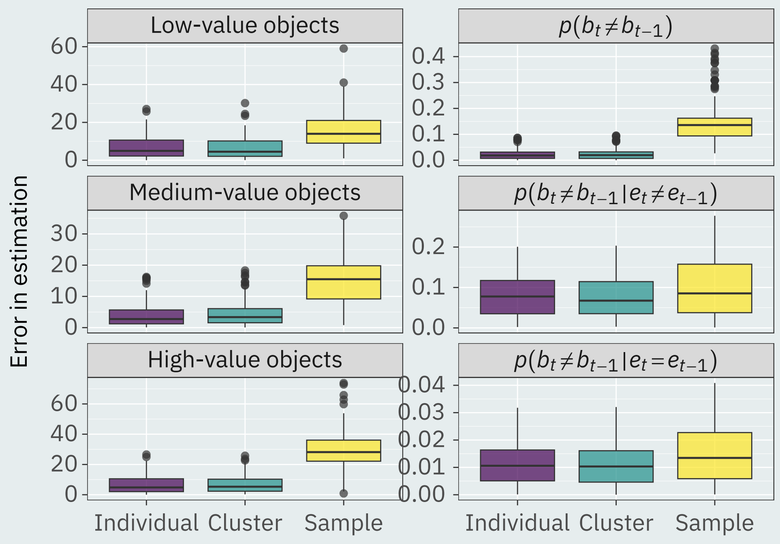
Lower Brier score and higher exp. mean log-likelihood $e^{\overline{LL}}$ mean better predictions:




Cluster-based DBNs perform as well as individual-specific DBNs and significantly better than the sample-level model (Friedman $p < 2.2 \times 10^{-16}$; post-hoc Wilcoxon: cluster vs. individual n.s., both vs. sample $r \approx 0.86$).

## 6. Key Updating Characteristics

Beyond overall accuracy we compare theoretically relevant characteristics: how often each belief state occurs (left) and how often the belief is updated (right) — overall and conditional on whether the sensory evidence changed.




The sample-level DBN is clearly off for *every* characteristic, whereas cluster-based DBNs stay on par with individual-specific ones.

## 7. Conclusion

(1) Cognitive processes such as belief updating differ between people yet share structure — subgroups are a principled compromise.
(2) In cluster-based models, individual differences are combined with the generalisability needed to predict previously unseen individuals.
(3) Cluster-based DBNs predict new trials about as accurately as individual-specific DBNs and clearly better than a single population-level model.

[1] University of Lübeck, [2] University of Hamburg