THE UNIVERSITY *of* EDINBURGH informatics

ilcc Institute for Language, Cognition and Computation

JRC for Logic

*Tsinghua University – University of Amsterdam Joint Research Centre for Logic*

# MCMC

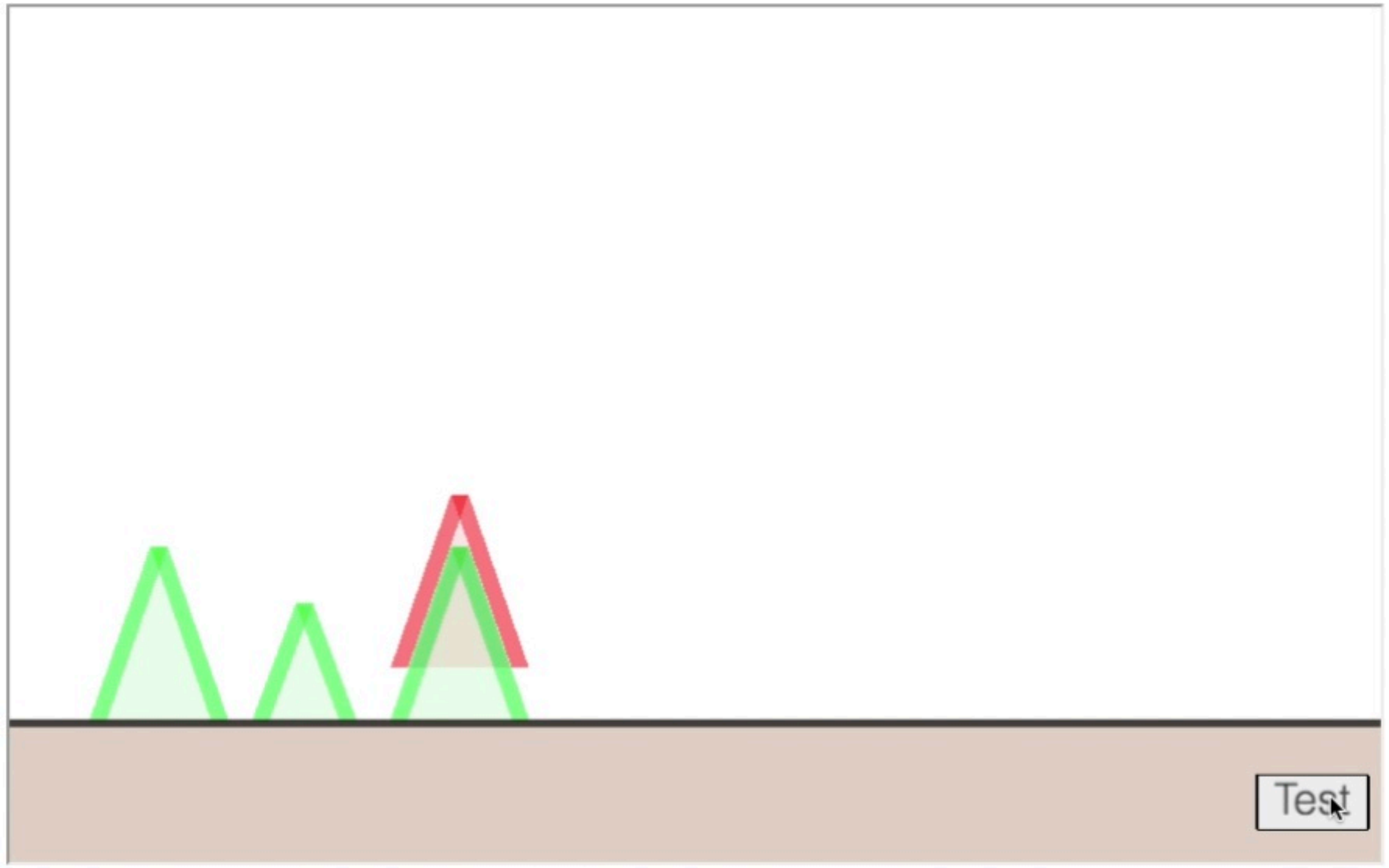
**Bonan Zhao (b.zhao@ed.ac.uk)**

Tsinghua Logic Summer School, July 2026


ZENDO

# Experiment

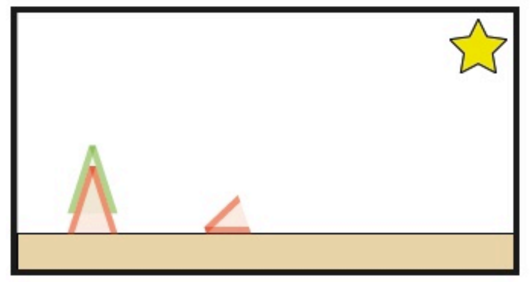
- **54 Children (8.9±1.1) & 50 Adult mTurkers**



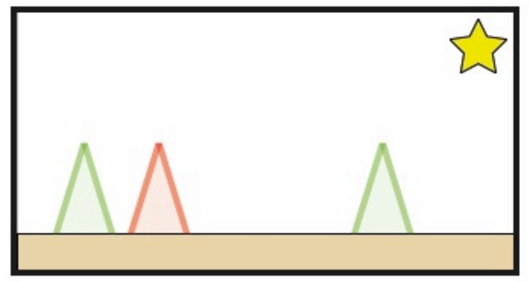

**Different tasks:**

Initial example

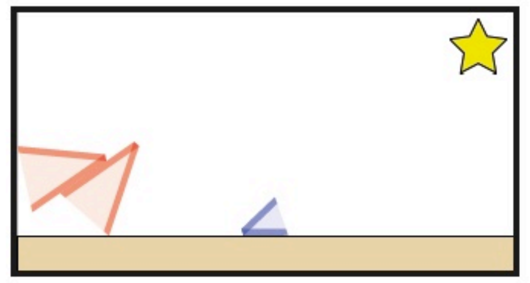
1. There is a red



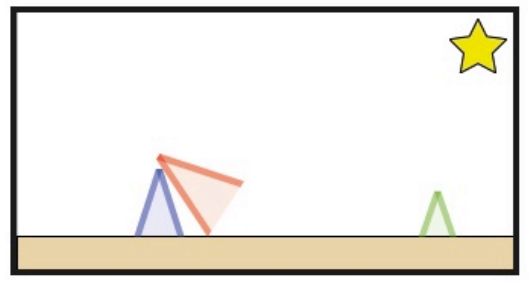
2. All same size



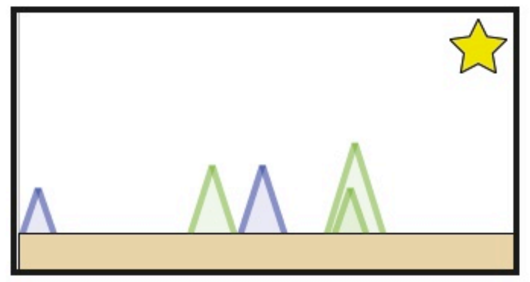
3. Nothing upright



4. Exactly 1 blue

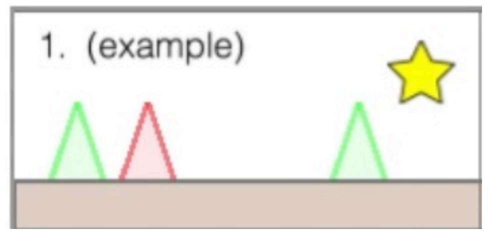


5. Something blue & small



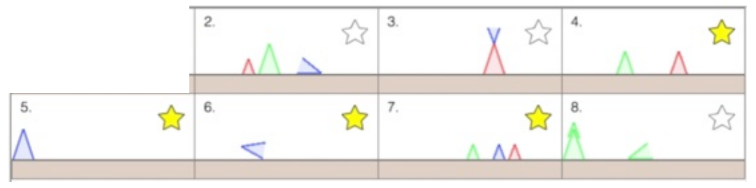
**Procedure:**

**Positive example** ×1



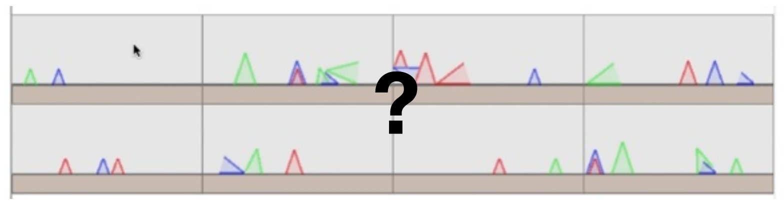

→

**Active learning** ×7



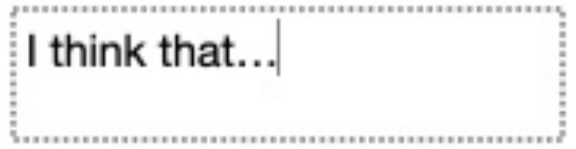
→

**Generalisation predictions** ×8




→

**Free guess** ×1



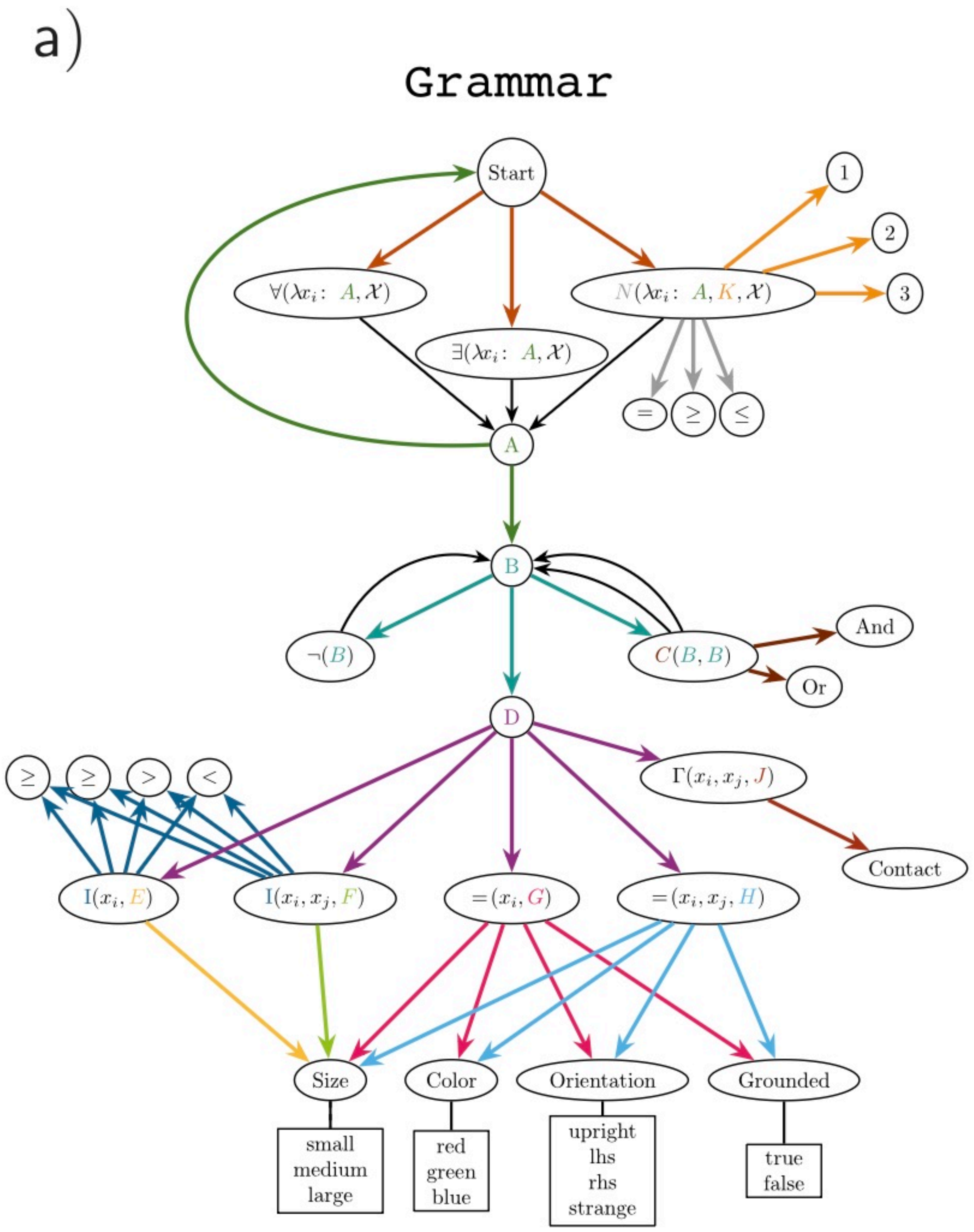
# Model

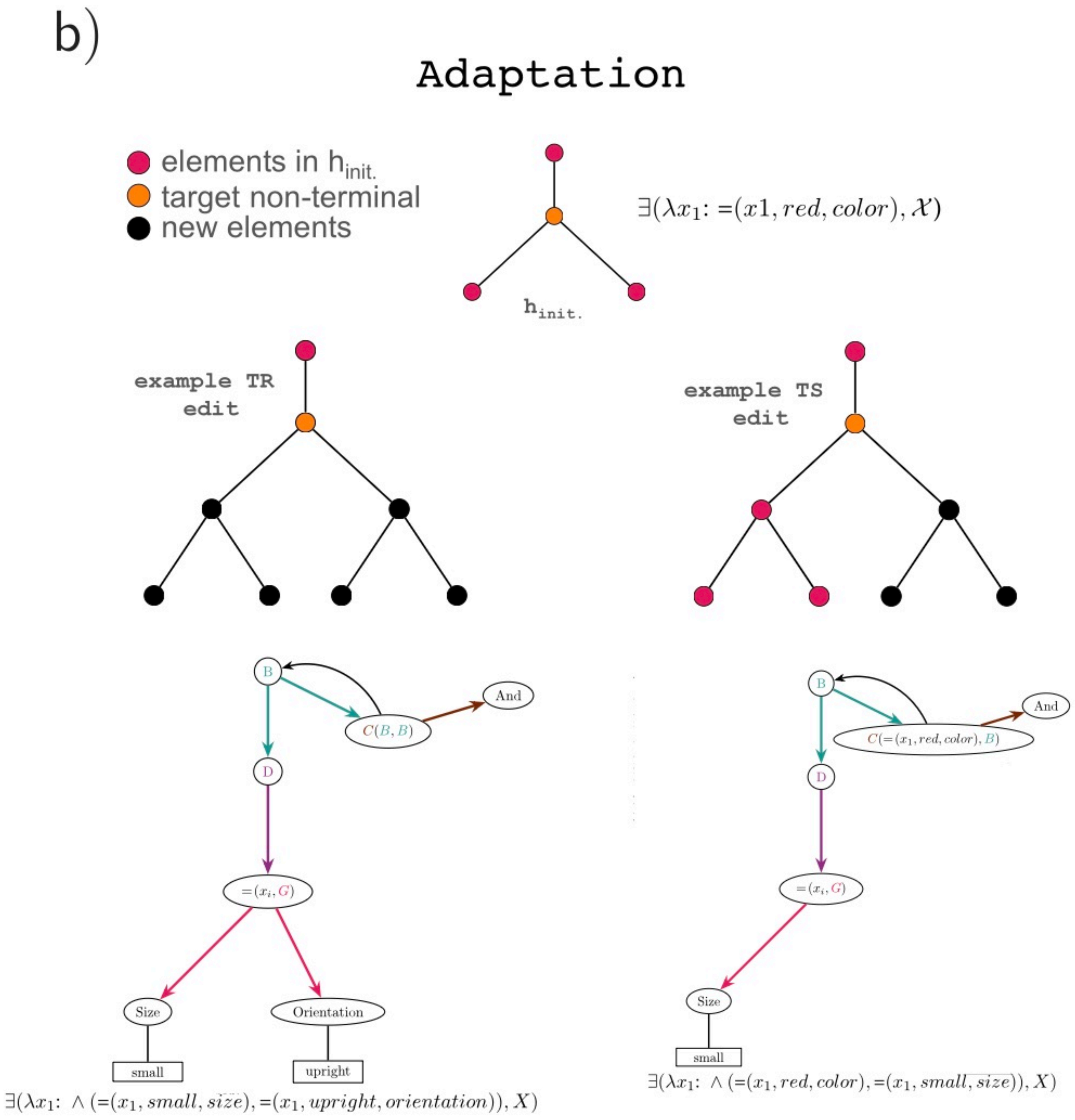
Concept Grammar.

| | |
|---|---|
| There exists an object $x_i$ such that... | $\exists(\lambda x_i :, \mathcal{X})$ |
| For all $x_i$ ... | $\forall(\lambda x_i :, \mathcal{X})$ |
| There exist(s) {at least, at most, exactly} $N$ objects $x_i$ such that... | $N_{\{<,>,=\}}(\lambda x_i :, N, \mathcal{X})$ |
| Feature $f$ of $x_i$ has value {larger, smaller, (or) equal} to $v$ | $\{<,>,\leq,\geq,=\}(x_i, v, f)$ |
| Feature $f$ of $x_i$ is {larger, smaller, (or) equal} to feature $f$ of $x_j$ | $\{<,>,\leq,\geq,=\}(x_i, x_j, f)$ |
| Relation $r$ between $x_i$ and $x_j$ holds | $\Gamma(x_i, x_j, r)$ |
| Booleans {and,or,not} | $\{\wedge, \vee, \neq\}(x)$ |

Note that $\{<,>,\geq,\leq\}$ comparisons only apply to the feature size.

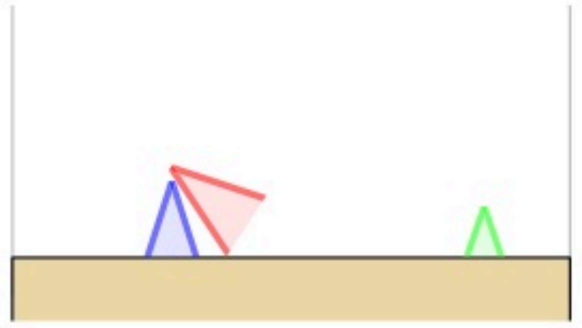
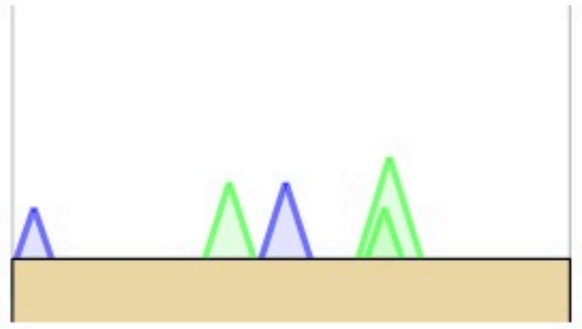
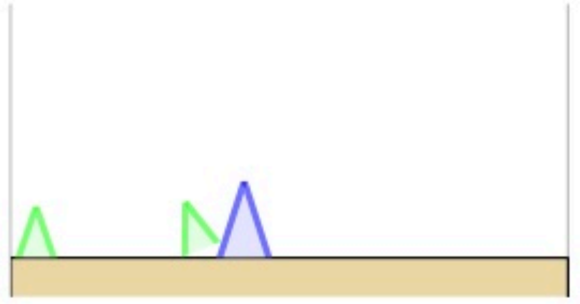
Fränken, J.-P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. Cognitive Psychology, 137, 101506.




Fränken, J.-P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. Cognitive Psychology, 137, 101506.

Five unknown test rules used during the main experiment.

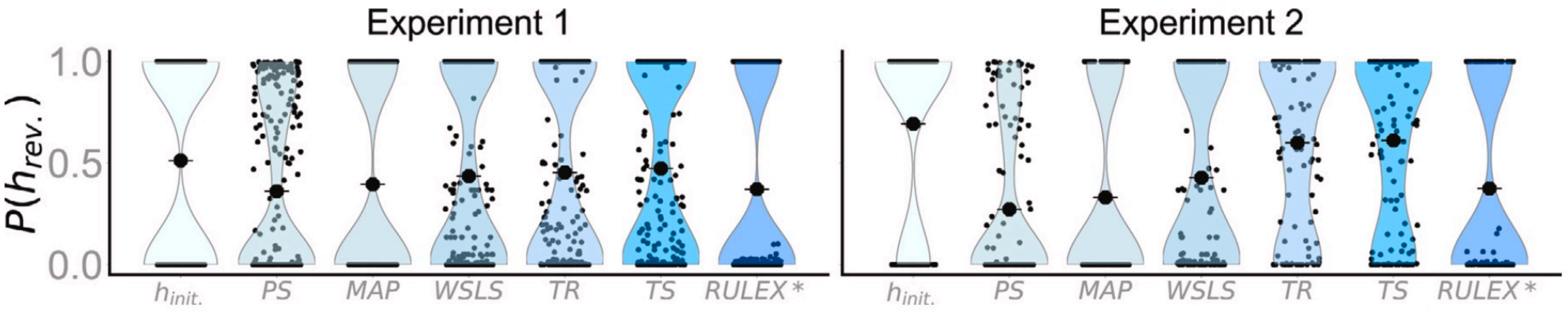
| Name | Rule | Prior Probability | Example |
|---|---|---|---|
| (1) Zeta | $\exists(\lambda x_1 : \; =(x_1, \text{red}, \text{color}), X)$ ("there is a red cone") | $\frac{\tau_{\text{color}}}{90}$ |  |
| (2) Iota | $N^{=}(\lambda x_1 : \; =(x_1, \text{blue}, \text{color}), 1, \mathcal{X})$ ("there is exactly one blue cone") | $\frac{\tau_{\text{color}}}{810}$ |  |
| (3) Upsilon | $\forall(\lambda x_1 : \; \neg(=(x_1, \text{upright}, \text{orientation})), X)$ ("no cone is upright") | $\frac{\tau_{\text{orientation}}}{270}$ |  |
| (4) Kappa | $\exists(\lambda x_1 : \; \wedge(=(x_1, \textit{small}, \text{size}), =(x_1, \text{blue}, \text{color}), X))$ ("there is a small blue cone") | $\frac{\tau_{\text{size}} \times \tau_{\text{color}}}{450}$ |  |
| (5) Omega | $\forall(\lambda x_1 : \; \vee(=(x_1, \textit{small}, \text{size}), =(x_1, \text{blue}, \text{color}), X))$ ("all cones are small or blue") | $\frac{\tau_{\text{size}} \times \tau_{\text{color}}}{450}$ |  |

Fränken, J.-P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. Cognitive Psychology, 137, 101506.

Best fitting model performances for participants' revised rule guesses ($h_{rev.}$).

| Experiment | Model | Mean $BIC_{h_{rev.}}$ | $N_{Best_{h_{rev.}}}$ | $b$ | $\lambda$ |
|---|---|---|---|---|---|
| Exp. 1 | $h_{init.}$ | 8.988 | 22.17 | – | – |
| | Posterior Sample (PS) | 9.817 | 6.00 | 5.0 | – |
| | Maximum a posteriori (MAP) | 11.164 | 4.17 | 7.0 | – |
| | Win–stay, lose–sample (WSLS) | 7.158 | 7.17 | 1.0 | – |
| | WSLS with Tree Regrowth (TR) | 6.368 | 12.17 | 3.0 | 10.0 |
| | **WSLS with Tree Surgery (TS)** | **5.917** | **15.17** | **5.0** | **10.0** |
| | RULEX* | 9.202 | 2.17 | – | 10.0 |
| Exp. 2 | $h_{init.}$ | 5.657 | 16.37 | – | – |
| | Posterior sample (PS) | 11.403 | 2.00 | 5.0 | – |
| | Maximum a posteriori (MAP) | 12.367 | 1.17 | 7.0 | – |
| | Win–stay, lose–sample (WSLS) | 6.573 | 0.37 | 1.0 | – |
| | WSLS with Tree Regrowth (TR) | 4.473 | 3.37 | 1.0 | 1.0 |
| | **WSLS with Tree Surgery (TS)** | **4.065** | **10.37** | **1.0** | **1.0** |
| | RULEX* | 9.220 | 2.37 | – | 3.0 |

# Predicting participants' hypotheses

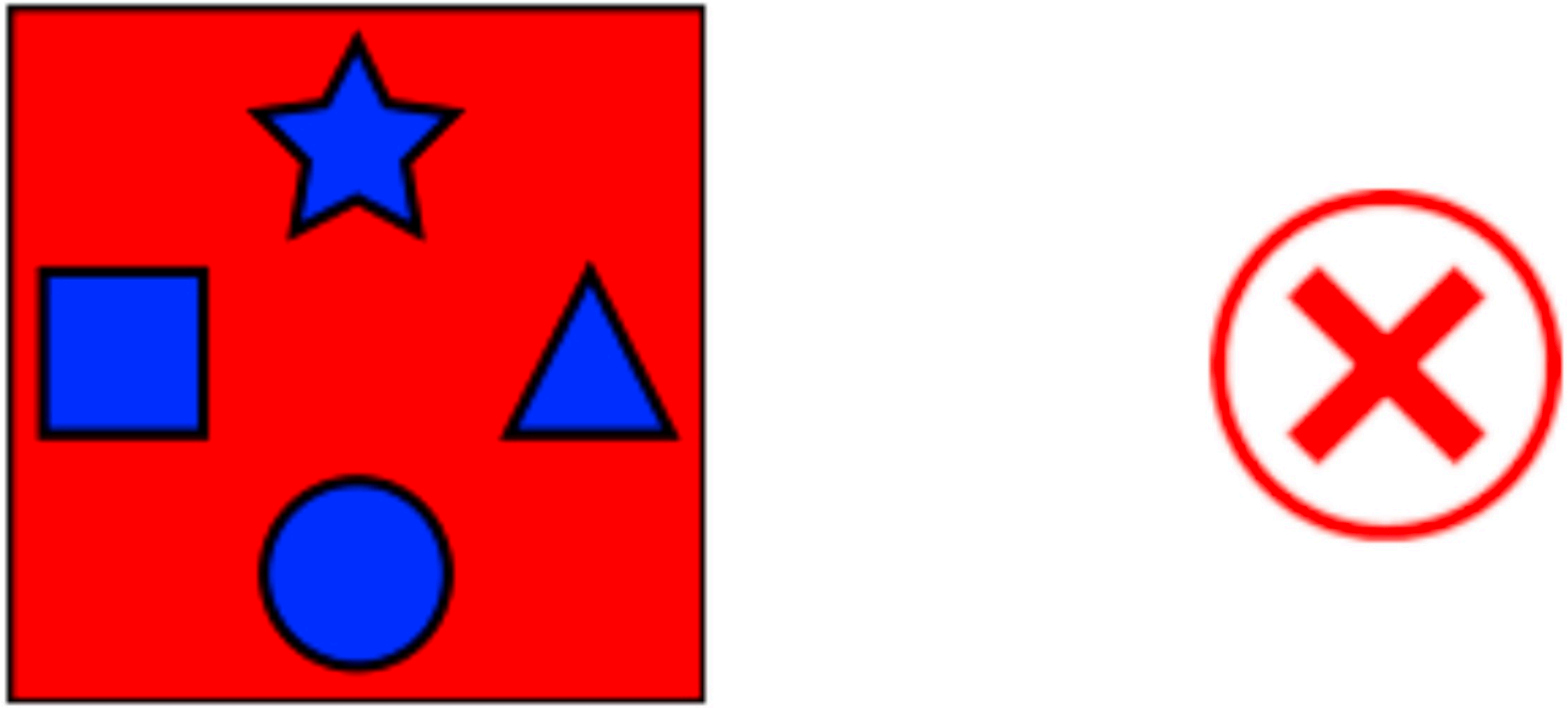



Fränken, J.-P., Theodoropoulos, N. C., & Bramley, N. R. (2022). Algorithms of adaptation in inductive inference. Cognitive Psychology, 137, 101506.

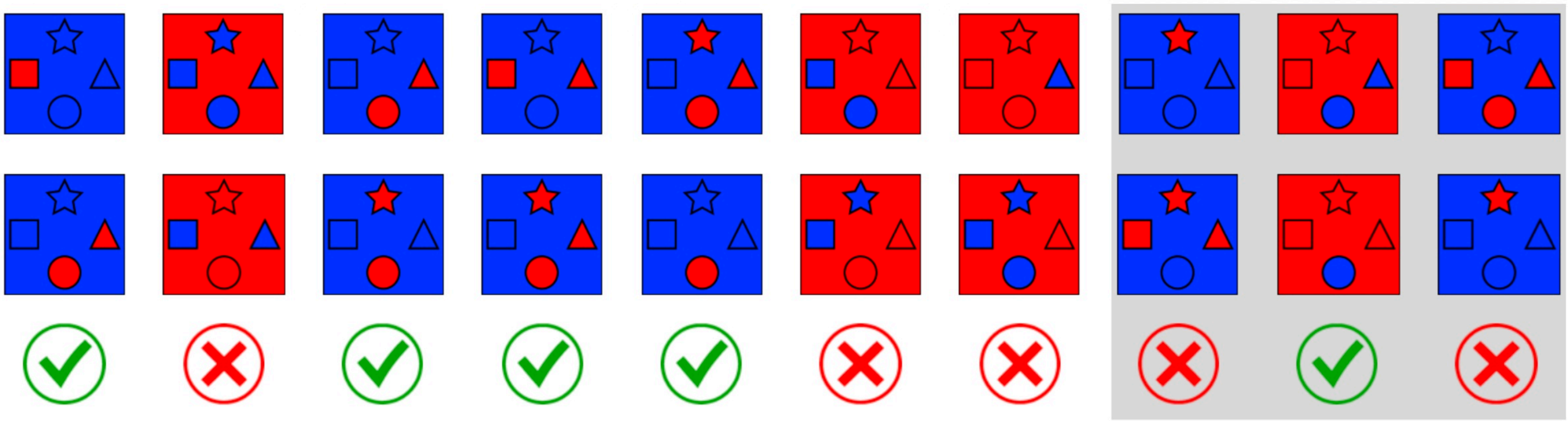
# Reasoning about complex causal relations

## With limited data

- People form rich and structured causal hypotheses.
- People update causal hypotheses locally.
- Local tree surgery provides a close account of human behavior.



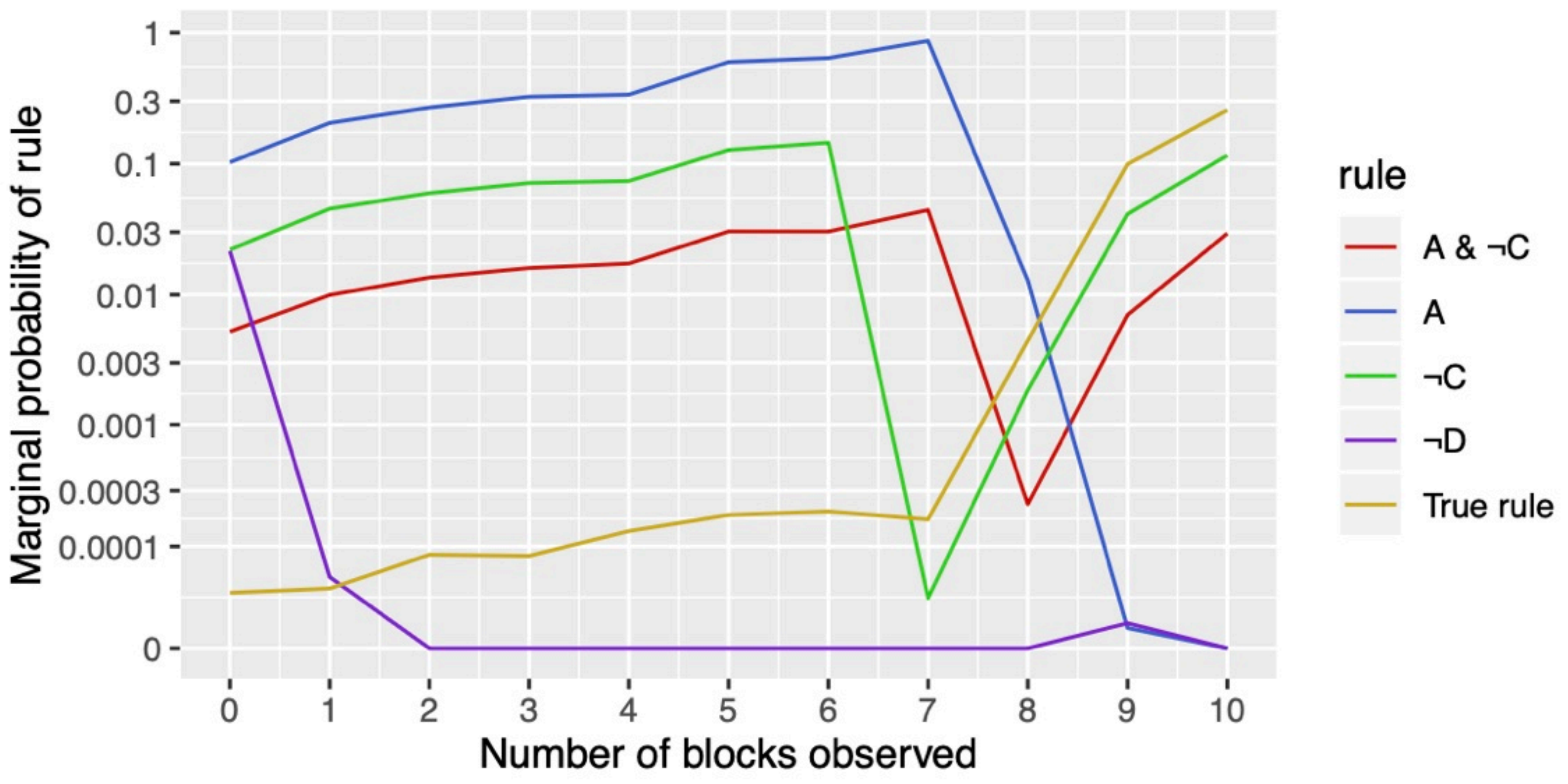
Distant, simpler true rule

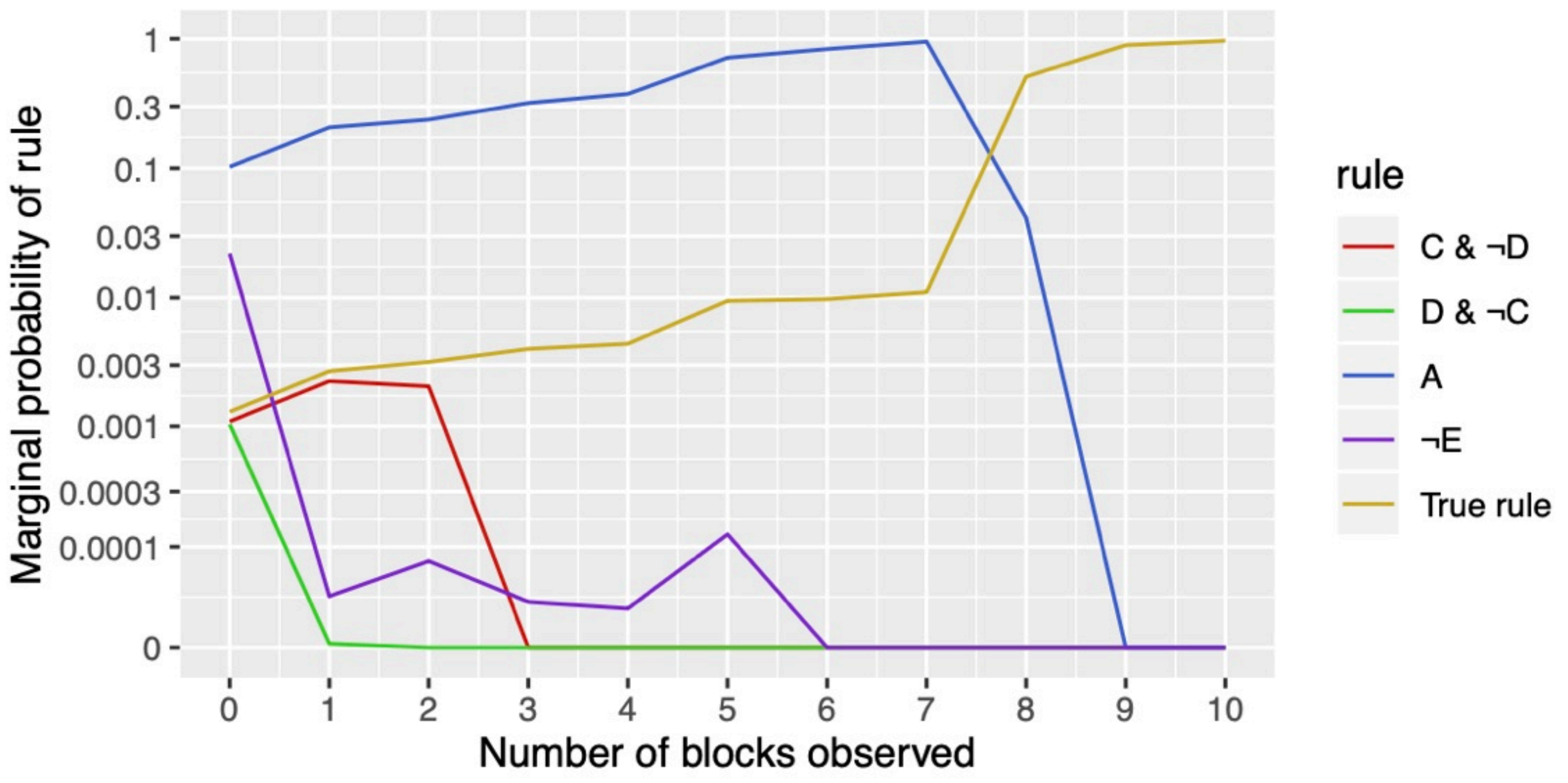


Near, longer true rule

Gelpi, R., Prystawski, B., Lucas, C. G., & Buchsbaum, D. (2020). Incremental hypothesis revision in causal reasoning across development. In Proceedings of the 42nd Annual Conference of the Cognitive Science Society. Cognitive Science Society.

# Model Predictions

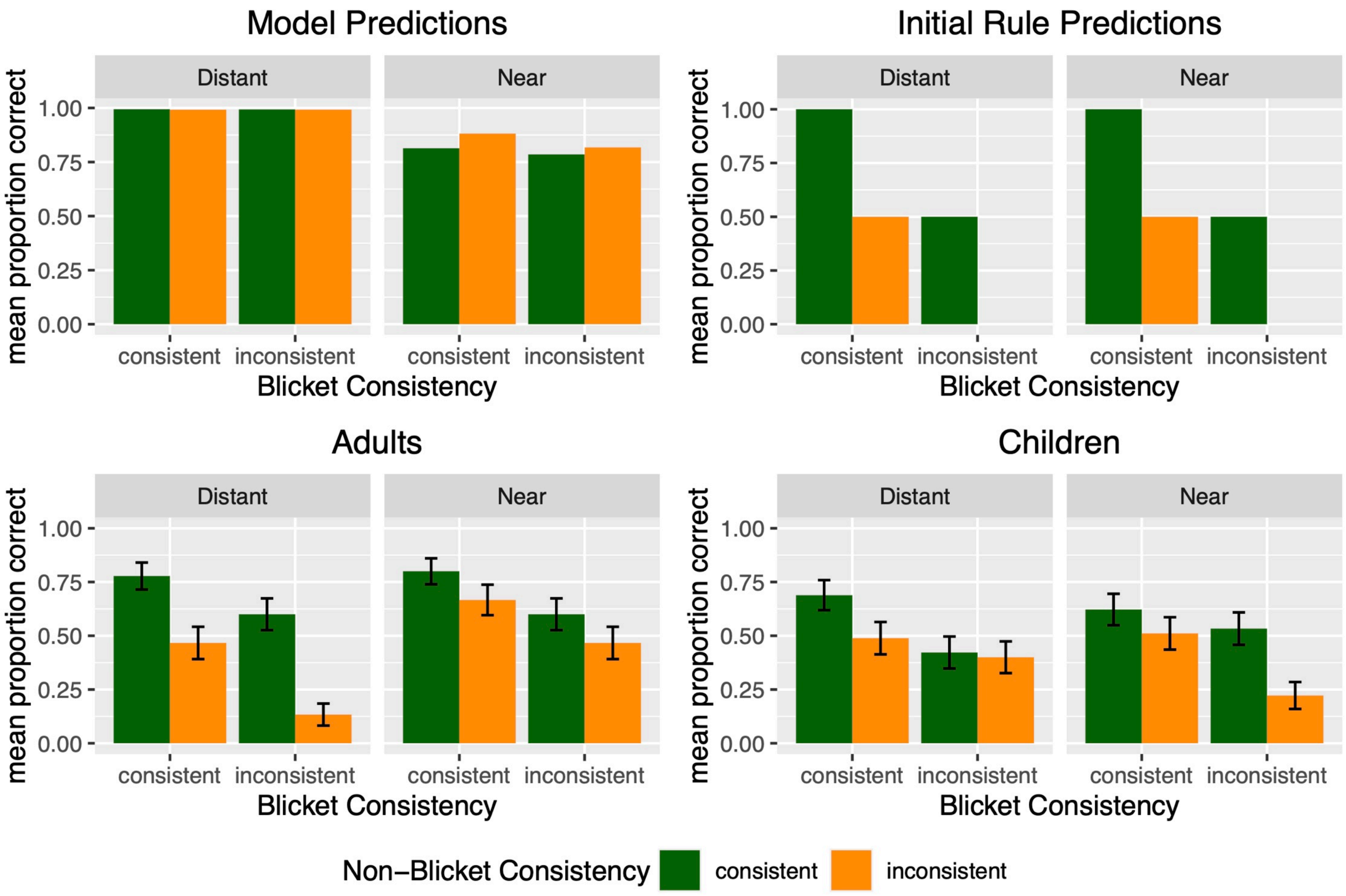




Gelpi, R., Prystawski, B., Lucas, C. G., & Buchsbaum, D. (2020). Incremental hypothesis revision in causal reasoning across development.
In Proceedings of the 42nd Annual Conference of the Cognitive Science Society. Cognitive Science Society.

# Main Result




Gelpi, R., Prystawski, B., Lucas, C. G., & Buchsbaum, D. (2020). Incremental hypothesis revision in causal reasoning across development. In Proceedings of the 42nd Annual Conference of the Cognitive Science Society. Cognitive Science Society.

# Reasoning about complex causal relations

## With limited data

- People form rich and structured causal hypotheses.
- People update causal hypotheses locally.
- Path-dependency.

# Marr's Levels of Analysis

**Computation**

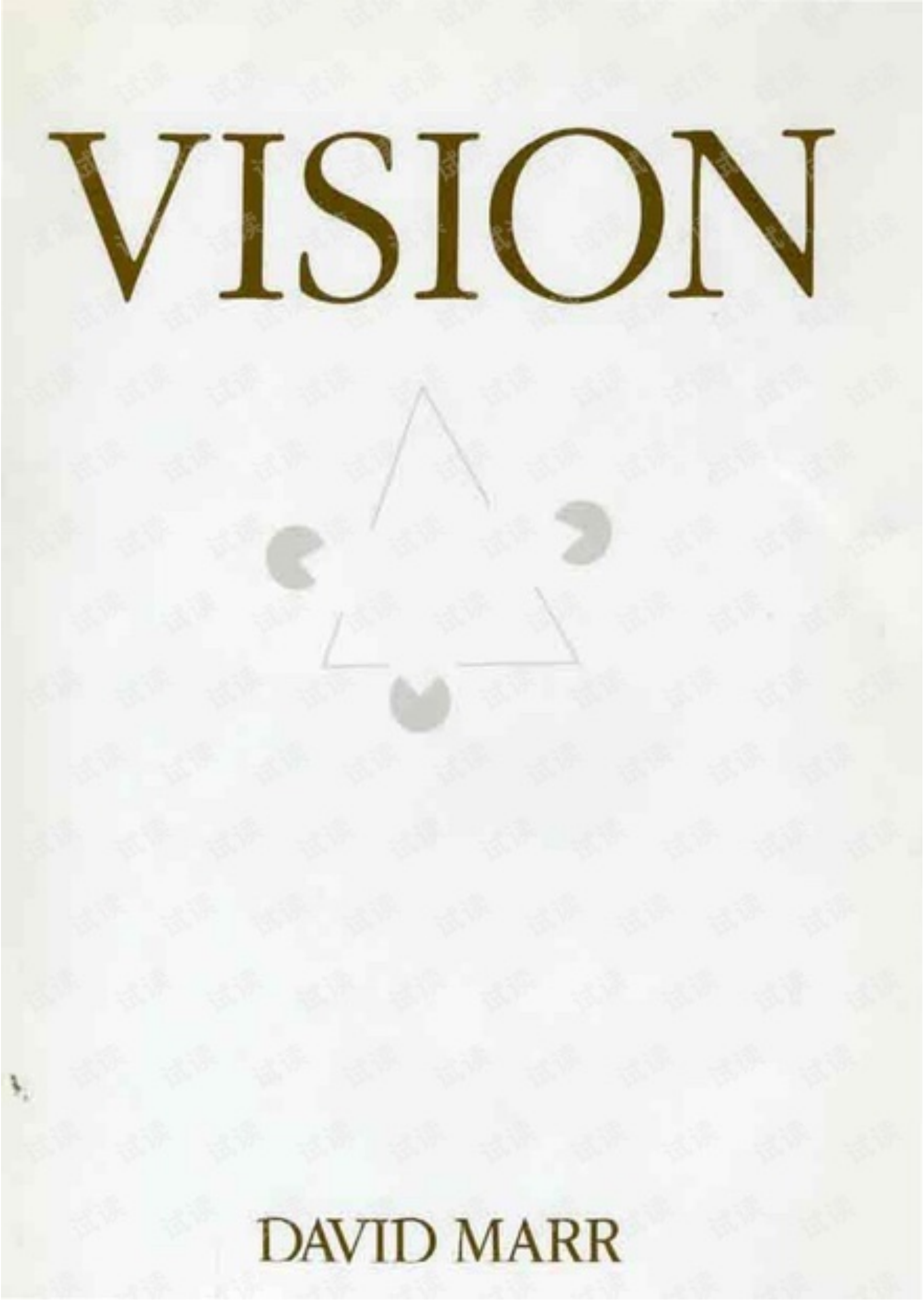
"What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?"

**Representation and algorithm**

"What is the representation for the input and output, and the algorithm for the transformation?"

**Implementation**

"How can the representation and algorithm be realized physically?"




Courtesy to Tom Griffiths, COS 360

# Marr's Levels of Analysis

**Computation**

"What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?"

$$P(h \mid D)$$

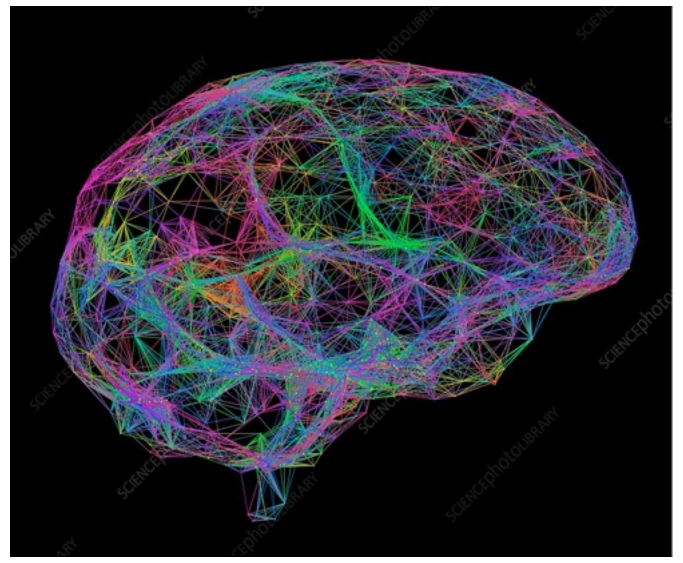
**Representation and algorithm**

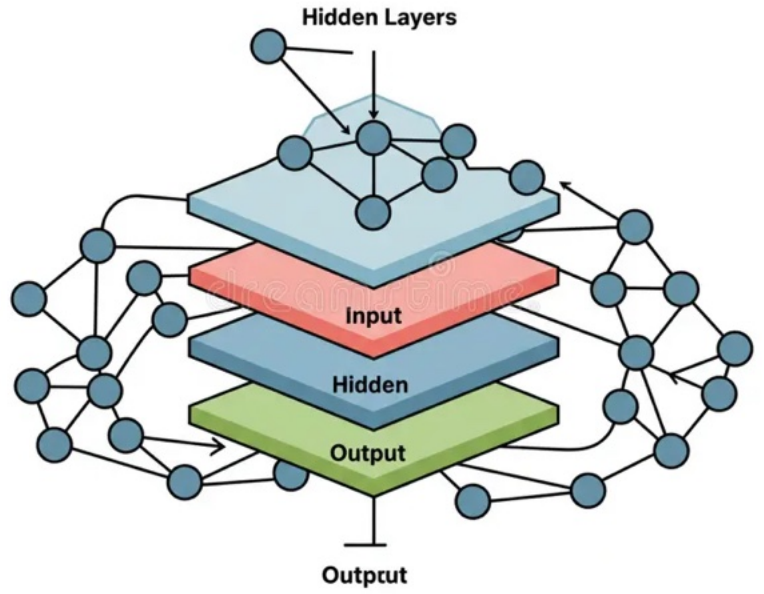
"What is the representation for the input and output, and the algorithm for the transformation?"

$$\hat{P}(h \mid D)$$

**Implementation**

"How can the representation and algorithm be realized physically?"






Courtesy to Tom Griffiths, COS 360

# Learning about causal models

## Or, the reverse inference of the true data generation function

- Causal model generates data
- Priors over possible causal models
- Using Bayes' rule to infer which causal model is the true one

# Learning about causal models

## Or, the reverse inference of the true data generation function

- Causal model generates data
- **Priors** over possible causal models
  - Evolutionary roots
  - Domain-specific knowledge
  - Open-endedness
- Using Bayes' rule to infer which causal model is the true one

# Learning about causal models

## Or, the reverse inference of the true data generation function

- Causal model generates data
- Priors over possible causal models
- Using Bayes' rule to **infer** which causal model is the true one
    - Approximation algorithms
    - Computational resource constraints

# Tomorrow

- Going back to what we can do with causal models