Investigating Relational Rule Learning

Helen Peng

What is Relational Reasoning?

Definition: A relation is a connection or link between two or more entities.
Example: Two objects are related if they have the same color.
Developmental Insights:

- Children's understanding shifts from concrete features to abstract relational similarities with age
- Higher-order relational similarity recognition emerges between ages 4 and 8
- Thought to be uniquely human, with limited capacity in non-human animals

Learning Relations

Why Is Learning Relations Hard:

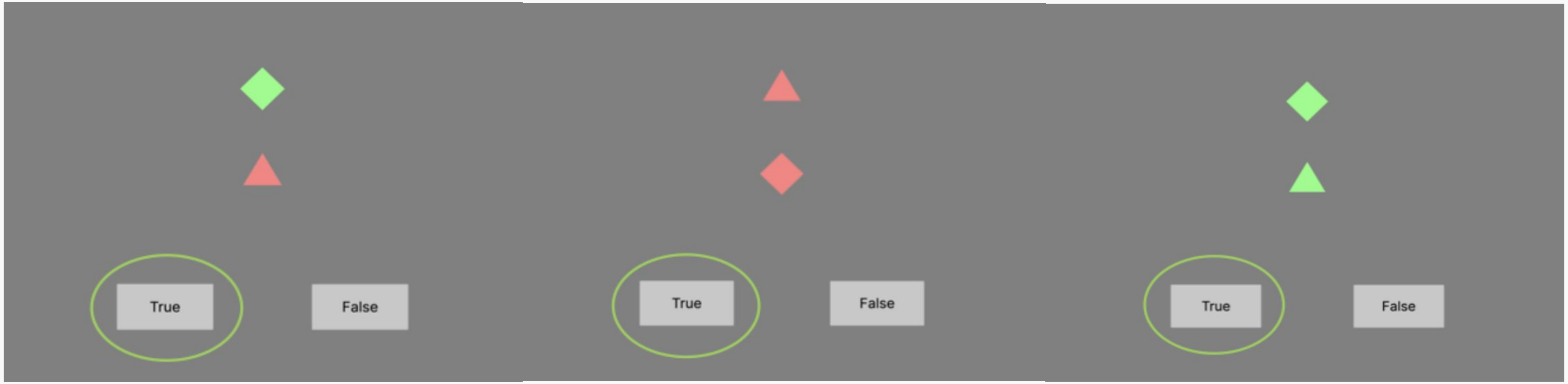
- Feature Variability: Objects often lack consistent surface features
- Learners may default to concrete traits instead of abstract relationships

Why It Matters:

- Promotes structural thinking beyond surface-level details
- Enables generalization of learned rules across novel contexts
- Critical for advanced reasoning in humans
- A key challenge in machine learning: enabling models to understand and apply relational patterns

Methods

The task consists of three compositional sub-rules: (1) when both shapes are green, the shape with more sides/angles must be on top; (2) when both shapes are red, the shape with more sides/angles must be on the bottom; and (3) in mixed conditions, the green shape must be on top and the red shape on the bottom.



01

How does the degree of understanding of a relational rule affect participants’ mastery rates across different task difficulty?

Criteria for Understanding

Complete/Full Understanding:

- Participant can fully state the relational rule in its entirety.

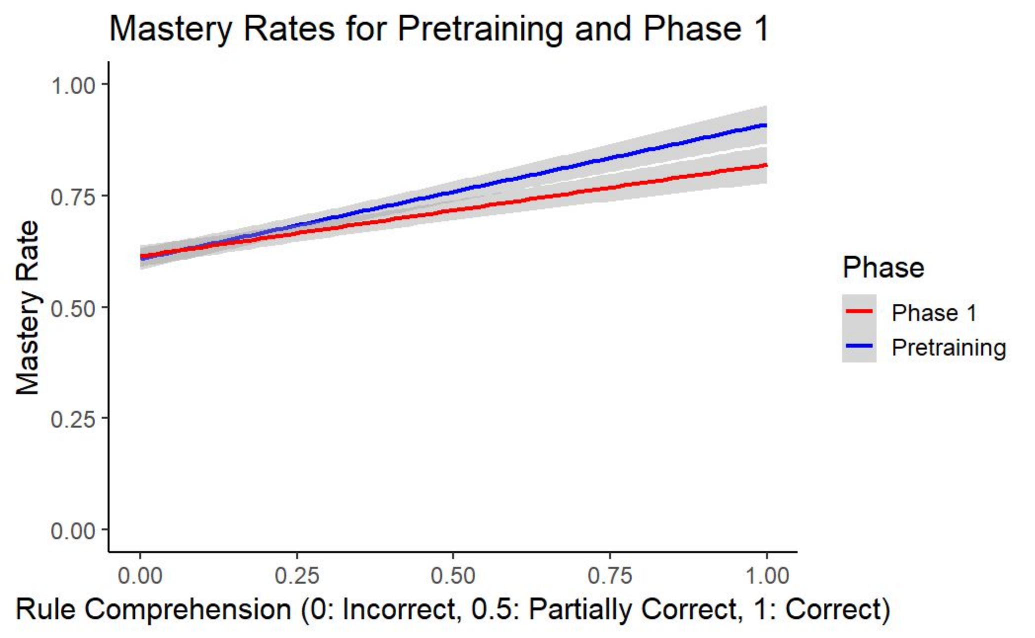
Partial Understanding:

- Participant mentions relevant features in the BothRed, BothGreen, and Mixed conditions, such as: (1) Sides or vertices in BothRed or BothGreen conditions. (2) Positional order (top/bottom) in the Mixed condition. In Phase 1, participant correctly identifies at least one of the rules.

Incomplete/No Understanding:

- Participant does not correctly identify any relevant relational features or rules.

T-tests reveal that mastery gaps between partial and incorrect understanding only become apparent as task difficulty increases.



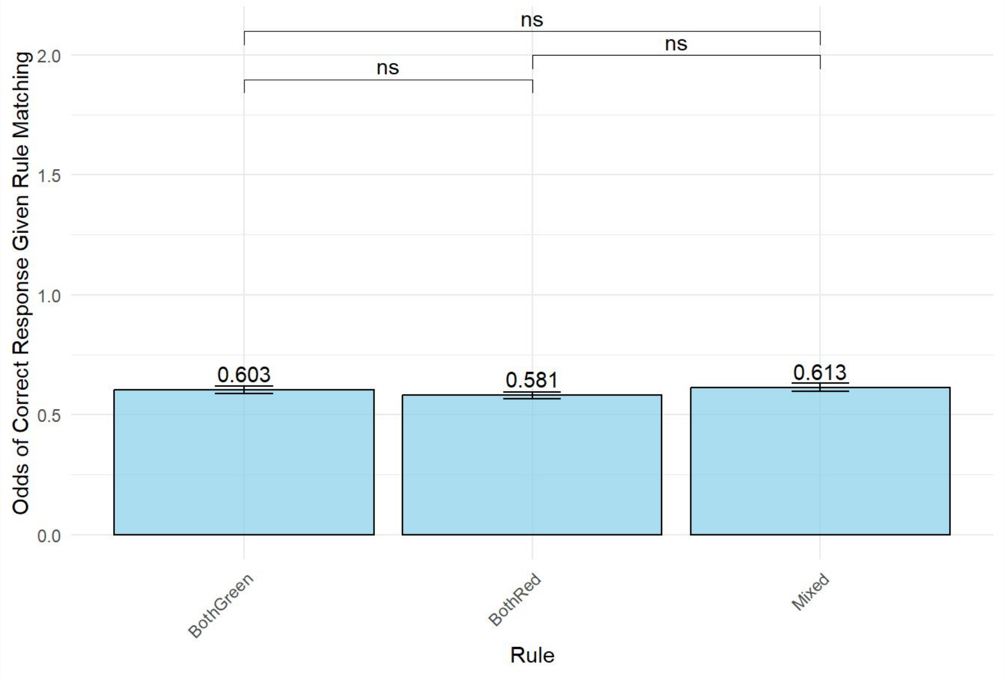
Pretraining Phase (Easier Task):
- Correct understanding led to higher mastery than partial or incorrect.
- Partial understanding showed no advantage over incorrect.

Phase 1 (Harder Task):
- Correct group outperformed both Partial and Incorrect groups.
- Partial group outperformed Incorrect group.

02

Does prior training on a specific relational rule improve task accuracy when the same rule is applied during testing?

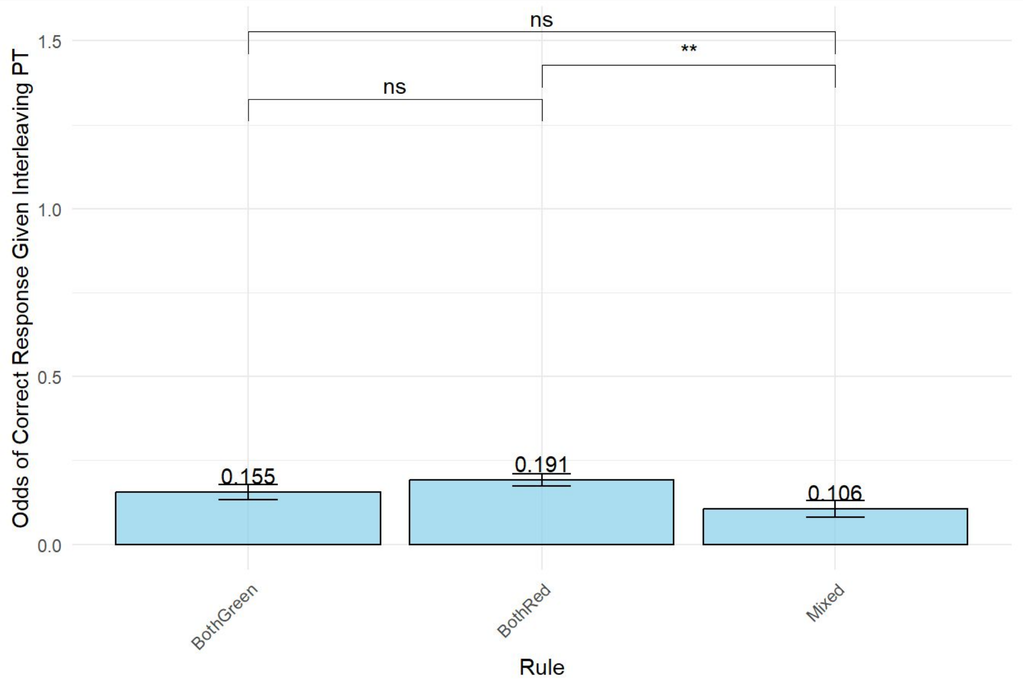
Multilevel modeling shows rule-matching improves accuracy, but the effect is modest and largely overshadowed by individual differences.



03

Does interleaved pre-training—mixing different relational rules during training—improve task accuracy during testing?

Multilevel modeling shows interleaving pre-training has no significant effect on accuracy, with individual differences driving most performance variation.



Discussion

Findings:

- Mastery gaps between partial and incorrect understanding appear only as tasks get harder.
- Rule-matching improves accuracy modestly; mostly individual differences.
- Interleaving pre-training has no significant effect; performance varies mainly by individual.

Implications:

- Focus on achieving full understanding early to handle complex tasks better.
- Tailor learning approaches to individual needs for greater effectiveness.

Next Steps:

- Investigate whether providing examples (learning support) improves rule mastery.
- Compare effects of interleaved versus blocked training.

Appendix

	RealCorrect		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.43	1.32 – 1.55	**<0.001**
MatchedRule	1.83	1.50 – 2.23	**<0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ \mathrm{Sub:Rule}}$	0.15		
$\tau_{00\ \mathrm{Sub}}$	0.15		
$\tau_{11\ \mathrm{Sub:Rule.MatchedRule}}$	0.54		
$\tau_{11\ \mathrm{Sub.MatchedRule}}$	0.37		
$\rho_{01\ \mathrm{Sub:Rule}}$	-0.33		
$\rho_{01\ \mathrm{Sub}}$	0.85		
ICC	0.15		
N_{Sub}	200		
N_{Rule}	3		
Observations	15118		
Marginal R^2 / Conditional R^2	0.017 / 0.165		

Level 1 (Within-Subject Model)

$$\mathrm{logit}(\pi_{ijr}) = \beta_{0jr} + \beta_{1jr} \cdot \mathrm{MatchedRule}_{ijr} + \epsilon_{ijr}$$

where:

- π_{ijr} is the probability of a correct response
- $\beta_{0jr} = 0.358$ is the subject-rule-specific intercept (baseline log-odds of a correct response)
- $\beta_{1jr} = 0.602$ is the subject-rule-specific slope for rule matching

Level 2 (Between-Subject and Between-Rule Models)

Models how the intercepts β_{0jr} and slopes β_{1jr} vary across subjects and subject-rule pairs

$$\beta_{0jr} = \gamma_{00} + u_{0j} + v_{0jr}$$

$$\beta_{1jr} = \gamma_{10} + u_{1j} + v_{1jr}$$

where:

- $\gamma_{00} = 0.358$ is the grand mean intercept
- $u_{0j} \sim N(0, \tau_{00_{Sub}} = 0.148)$ is the variability of subject means around the grand mean intercept
- $v_{0jr} \sim N(0, \tau_{00Sub:Rule} = 0.1511)$ is the subject-rule-level deviation
- $\gamma_{10} = 0.602$ is the grand mean slope
- $u_{1j} \sim N(0, \tau_{11_{Sub.MatchedRule}} = .16)$ is the subject-level deviation for slope
- $v_{1jr} \sim N(0, \tau_{11Sub:Rule.MatchedRule} = .75)$ is the subject-rule-level deviation for slope

Appendix

	RealCorrect		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.57	1.43 – 1.73	**<0.001**
InterleavedPTTRUE	1.16	0.90 – 1.49	0.248
Random Effects			
σ^2	3.29		
$\tau_{00\ Sub:Rule}$	0.30		
$\tau_{00\ Sub}$	0.16		
$\tau_{11\ Sub:Rule.InterleavedPTTRUE}$	0.47		
$\tau_{11\ Sub.InterleavedPTTRUE}$	0.54		
$\rho_{01\ Sub:Rule}$	-0.69		
$\rho_{01\ Sub}$	-0.39		
ICC	0.14		
N_{Sub}	200		
N_{Rule}	3		
Observations	15118		
Marginal R^2 / Conditional R^2	0.001 / 0.137		

Level 1 (Within-Subject Model)

$$\mathrm{logit}(\pi_{ijr}) = \beta_{0jr} + \beta_{1jr} \cdot \mathrm{InterleavedPTTRUE}_{ijr} + \epsilon_{ijr}$$

where:

- π_{ijr} is the probability of a correct response
- $\beta_{0jr} = 0.45186$ is the subject-rule-specific intercept (baseline log-odds of a correct response)
- $\beta_{1jr} = 0.14853$ is the subject-rule-specific slope for interleaving pretraining

Level 2 (Between-Subject and Between-Rule Models)

Models how the intercepts β_{0jr} and slopes β_{1jr} vary across subjects and subject-rule pairs

$$\beta_{0jr} = \gamma_{00} + u_{0j} + v_{0jr}$$

$$\beta_{1jr} = \gamma_{10} + u_{1j} + v_{1jr}$$

where:

- $\gamma_{00} = 0.45186$ is the grand mean intercept
- $u_{0j} \sim N(0, \tau_{00_{Sub}} = 0.16)$ is the variability of subject means around the grand mean intercept
- $v_{0jr} \sim N(0, \tau_{00Sub:Rule} = 0.30)$ is the subject-rule-level deviation
- $\gamma_{10} = 0.14853$ is the grand mean slope
- $u_{1j} \sim N(0, \tau_{11_{Sub.InterleavedPTTRUE}} = 0.54)$ is the subject-level deviation for slope
- $v_{1jr} \sim N(0, \tau_{11Sub:Rule.InterleavedPTTRUE} = 0.47)$ is the subject-rule-level deviation for slope

References

Kittur, A., Hummel, J. E., & Holyoak, K. J. (n.d.). *Feature- vs. Relation-Defined Categories: Probab(alistic)ly Not the Same*.

Kotovsky, L., & Gentner, D. (1996). *Comparison and Categorization in the Development of Relational Similarity*.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin’s mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(2), 109–130. https://doi.org/10.1017/S0140525X08003543