



# Dissociating Explicit from Implicit Strategies In Two-Dimensional Category Learning



Northwestern University

Ben Reuveni

Catherine Ives-Louter

Paul J. Reber

Brain Imaging & Memory Lab

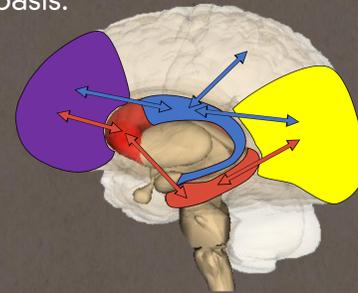
## Introduction

Human category learning is driven two qualitatively distinct types of mechanisms:

- Explicit** - Rule (or exemplar) based:
    - Strategies rely on explicit reasoning processes consisting of one or more verbalizable rules to learn the category.
    - Pre-Frontal Cortex  $\leftrightarrow$  Medial Temporal Lobe
  - Implicit** - Information Integration:
    - An associative mechanism that integrates information (II) across stimulus dimensions learned through dopamine dependent reinforcement, and happens outside awareness.
    - Non-MTL, Basal Ganglia important, plus ?
- We are interested in implicit learning in the context of intuition.
- How does implicit information flow to the PFC for decision making?

## PINNACLE

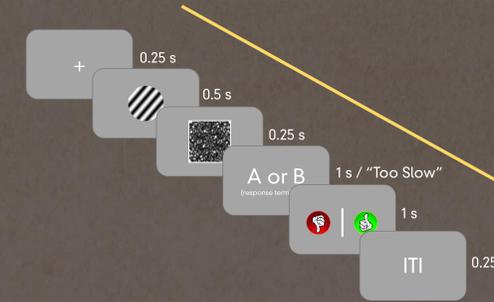
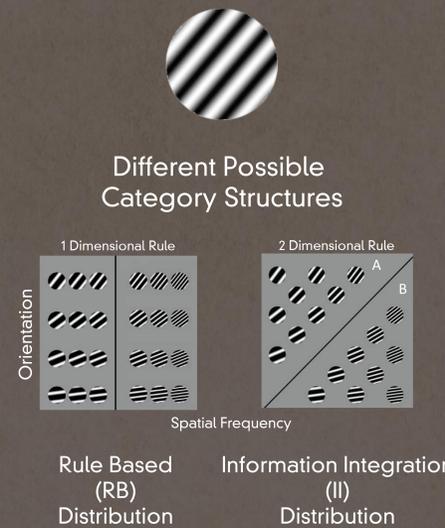
The "Parallel Interactive Neural Networks Active in Category Learning" model incorporates separate, parallel and independent **explicit** and **implicit** categorization mechanisms that compete to provide categorization judgments about visual stimuli on a trial-by-trial basis.



## Methods

## Feedback-Based trial and error paradigm

Is this an "A" or a "B" ?



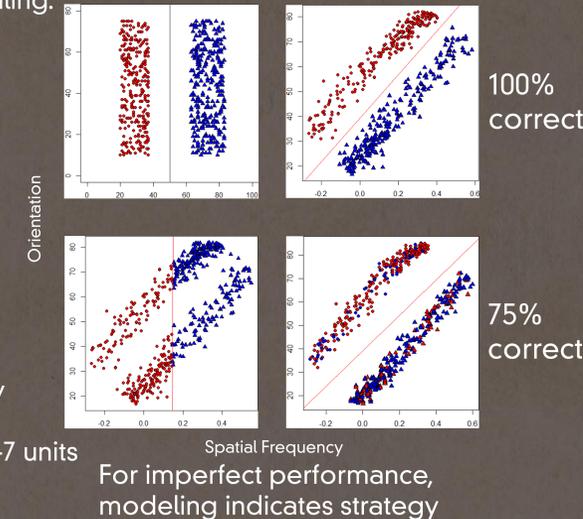
n = 122 undergraduate participants for course credit

11 "II" conditions:

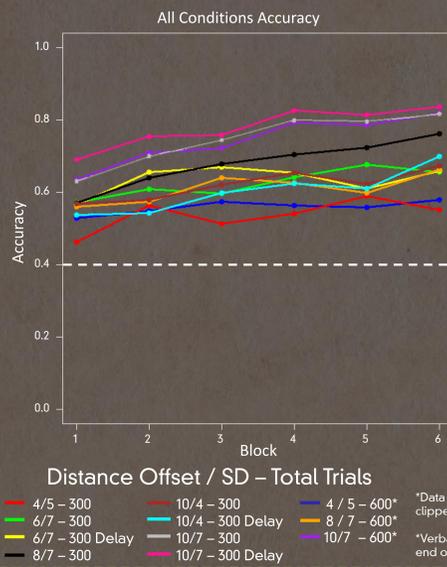
- Categories separated by 2d boundary
- Each category cigar shaped
- SD within category ranged between 4-7 units
- Category distance was 1.2 to 1.4 SD

## Modeling

Decision Bound Theory (DBT) provides a first step towards dissociating strategies in cases where performance is above chance but below ceiling.



## Results

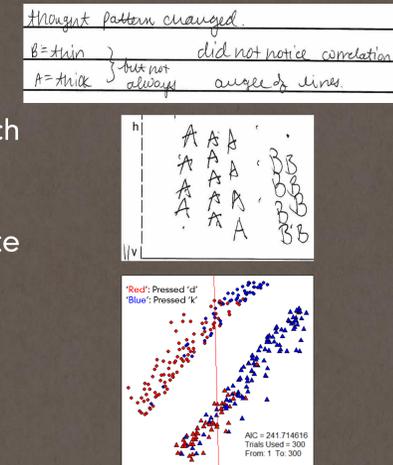


After completion, participants are interviewed about their conscious knowledge of the II category. The interview starts with an open-ended, minimally-guided, free response probe as to how the participant solved the task. Leading words like "strategy," "rules," and "patterns" are avoided. The two-dimensional space of frequency and orientation is described, and participants attempt to diagram the category structure. 4 independent raters then codified the interview responses into two categories with inter-rater reliability = 90%:

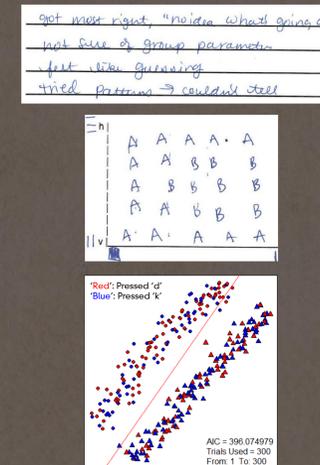
- Can verbalize a 1D rule
- Can verbalize a 2D rule

- Please tell us about the experiment.
- Consider this space. For each point imagine the corresponding sine-wave. If you believe it is an 'A', write "A". If you believe it is a 'B', write "B".
- DBT Modeling with 2 degrees of freedom.

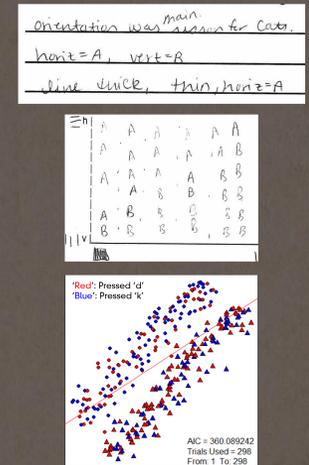
## Example of a RB Strategy



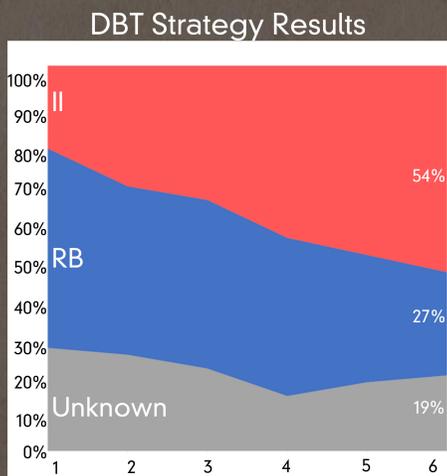
## Example of an Implicit II Strategy



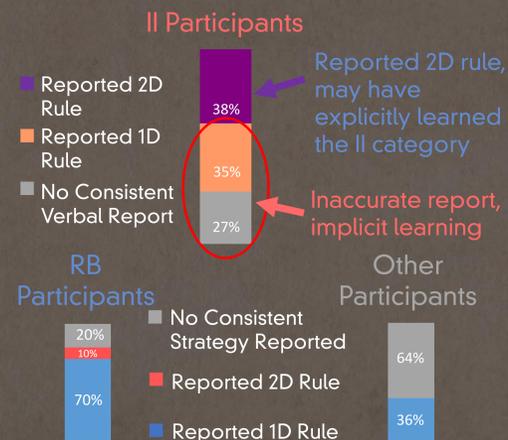
## Example of an Explicit II Strategy



## Strategy



## Verbal Report Results



## Discussion

Current best practices rely on category structure and data modeling to infer strategy.

When investigating the neural mechanisms of learning during an "II" task, model-fitting does not fully distinguish between participants using an **explicit** versus **implicit** II strategy.

The inclusion of structured, minimal-intervention, free-response post-experimental interviews assesses verbalizability across participants.

Within the PINNACLE framework:

- Verbal report reflects knowledge in the MTL-PFC system
- Explicit-II likely depends on the MTL-PFC (RB)
- Only Implicit-II reflects non-MTL memory systems

This subset of 'explicit II' strategy users indicates that prior dissociations between category systems underestimates category learning differences.

## Future Directions & References

- By incorporating assessments of conscious knowledge, future work can more accurately identify specific characteristics of **implicit** visual category learning by excluding these **explicitly**-driven participants.
- Utilizing PINNACLE and the verbal interview presented here, we plan to image memory systems interactions in the context of visual category learning. The ability to dissociate **implicit** from **explicit** II strategies will be critical to our analysis.

Nomura, E. M., & Reber, P. J. (2012). Combining computational modeling and neuroimaging to examine multiple category learning systems in the brain. *Brain sciences*, 2 (2), 174-202.

Ives-Louter, C., Reuveni, B., & Reber, P. J. (2016, April). Using a novel category learning paradigm to reliably produce implicit learning. Poster presented at the annual meeting of the Cognitive Neuroscience Society, New York, NY. [F171 - 4/5/16 - Americas Hall II]

