The representation and retrieval of general versus specific category knowledge

Marlie C. Tandoc, Sarah H. Solomon, Jacob A. Parker, Alex Gordienko, Anna C. Schapiro

University of Pennsylvania, Department of Psychology

Penn Computational Cognitive Neuroscience Lab

Introduction

Building useful knowledge requires learning both **general** and **specific** information (most dogs bark, but your friend's dog bites)

There is evidence that different neural systems may be dedicated towards learning this information[1-4]

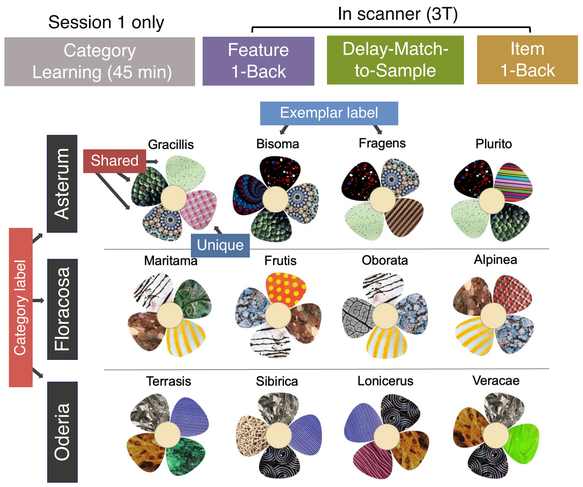
How does the brain **represent** and flexibly **retrieve** general versus specific knowledge?

How is this knowledge learned **rapidly** and how might it transform over **time**?

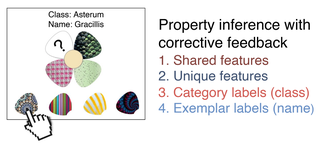
Design

n = 30, two-session fMRI (1-2 weeks apart)

Categories with **general** (shared features, category labels) and **specific** (unique features, exemplar labels) information

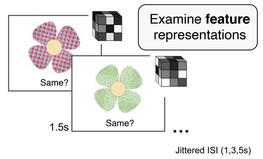


Category Learning

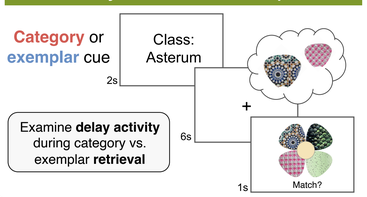


Property inference with corrective feedback
1. Shared features
2. Unique features
3. Category labels (class)
4. Exemplar labels (name)

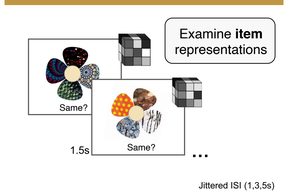
Feature 1-back



Delay-Match-to-Sample

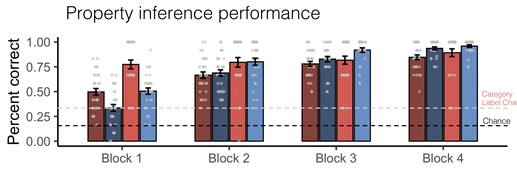


Item 1-Back

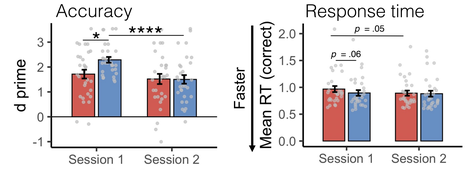


Behavioral Results

Category Learning



Delay-Match-to-Sample



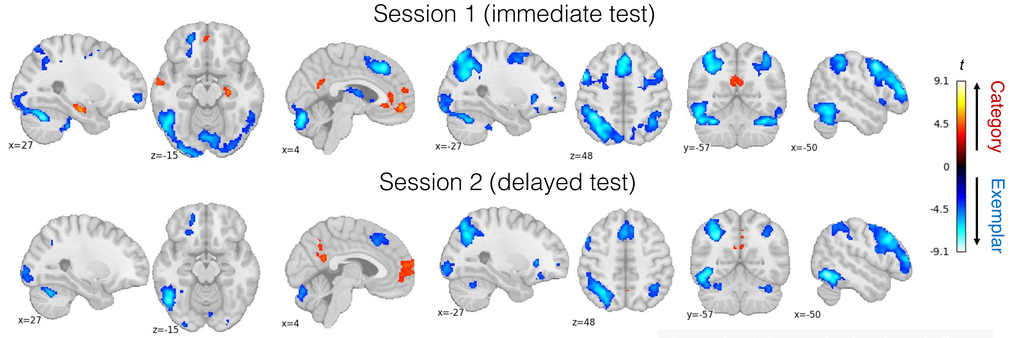
Shared Feature | Unique Feature | Category Label | Exemplar Label

Rapid learning of **general** and **specific** information. Memory for **exemplars** gets worse with time while **categories** persists

Retrieval Results

Delay-Match-to-Sample

Whole-brain univariate analysis: Category — exemplar retrieval



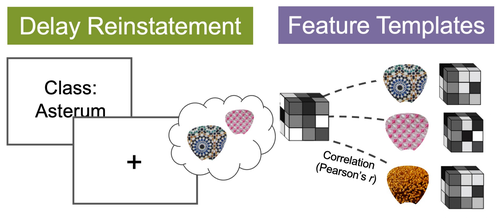
Clusters survived multiple comparisons using non-parametric permutation testing at $p < .05$
Analyses restricted to the first 3 seconds of the delay period on correct trials only

In session 2, less parietal and cerebellum for **exemplars** and less dorsal striatum for **categories**

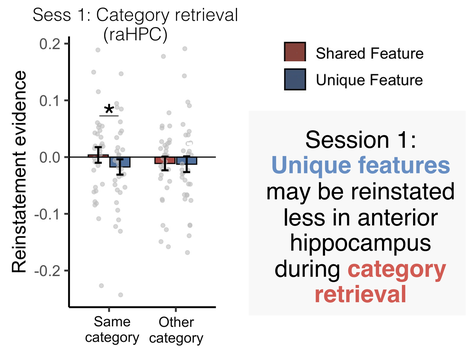
Category retrieval: Medial prefrontal cortex, anterior hippocampus, anterior temporal lobe, precuneus

Exemplar retrieval: Visual cortex, parietal cortex, dorsal and lateral prefrontal cortex, cerebellum

Preliminary hippocampal reinstatement analysis



Reinstatement Evidence
Correlation between delay period (retrieval) and feature patterns

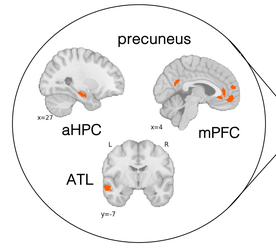


Session 1: **Unique features** may be reinstated less in anterior hippocampus during **category retrieval**

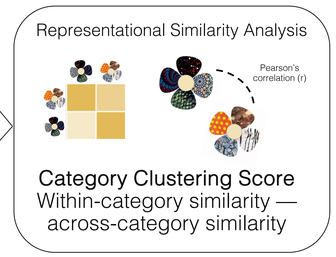
Item Representations

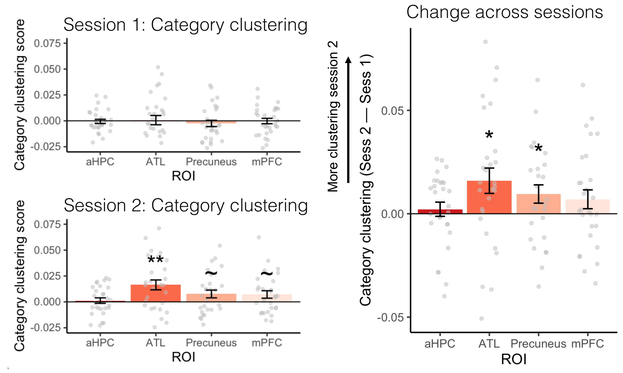
Item 1-Back

Category retrieval ROIs from delay-match-to-sample



Look for category clustering of items in these ROIs





Items show **category clustering** in category retrieval regions, but only after 1-2 weeks

Conclusions

Distinct brain regions support the retrieval of general vs. specific knowledge, consistent with work showing some of these regions employ representations at different levels of abstraction[2]

Representations of **items from the same category become more overlapping** in regions important for retrieving categories, but **only after consolidation**

Next Steps

Feature RSA, hippocampal subfield analyses, searchlight, relate neural measures to behavior, examine the content of reinstatement during category vs. exemplar retrieval

tandoc@sas.upenn.edu

References

1. Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1711), 20160049.
2. Bowman, C. R., Iwashita, T., & Zeithamova, D. (2020). Tracking prototype and exemplar representations in the brain across learning. *elife, 9*, e59360.
3. Samborska, V., Butler, J. L., Walton, M. E., Behrens, T. E., & Akam, T. (2022). Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. *Nature Neuroscience, 25*(10), 1314-1326.
4. McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419.