

# How surprising information shifts our attention and influences learning

Em Heffernan

## **Category learning:** grouping items by attending to shared features

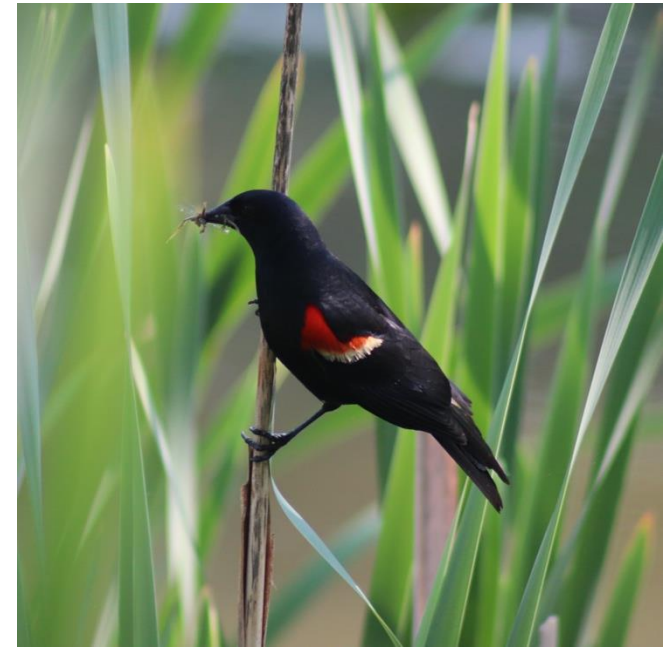
Category learning is an imperfect process. Because they look similar, we might assume that the center bird is related to the bird on the left, but it's actually a blackbird!

**savannah sparrow**

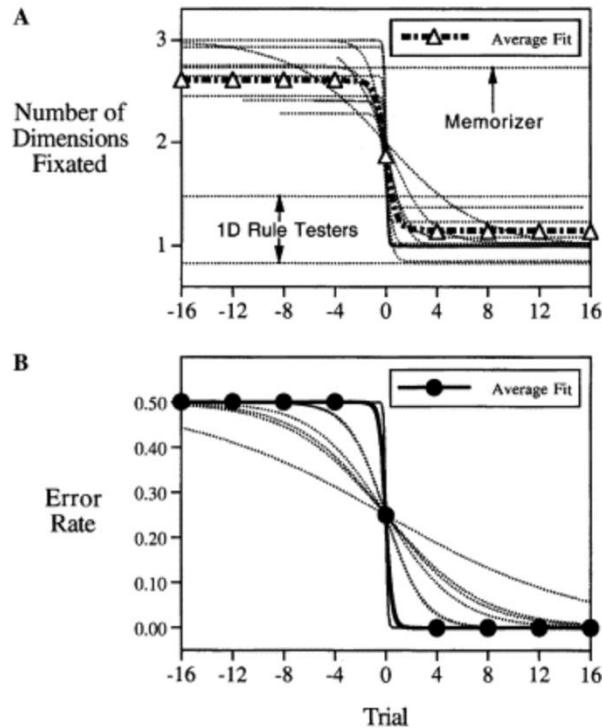


**another sparrow??**  
**female red-winged black bird**

**red-winged black bird**



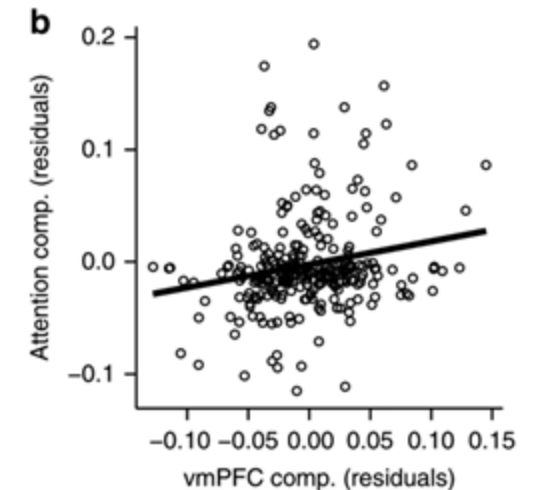
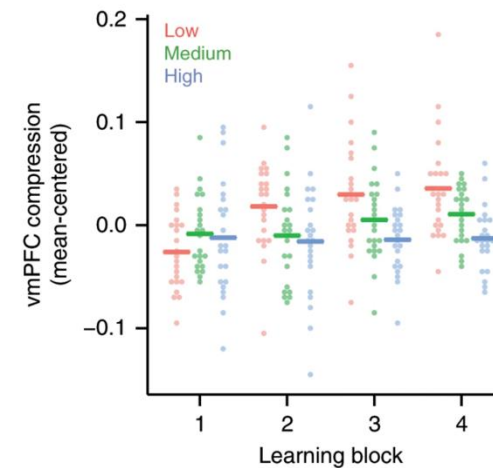
# Attentional tuning in category learning



(Rehder & Hoffman, 2005)

We can measure attention during learning via eye or mouse-tracking

(Rehder & Hoffman, 2005; Blair et al., 2009; Zaki & Salmi, 2009; \*Chen et al., 2013...)



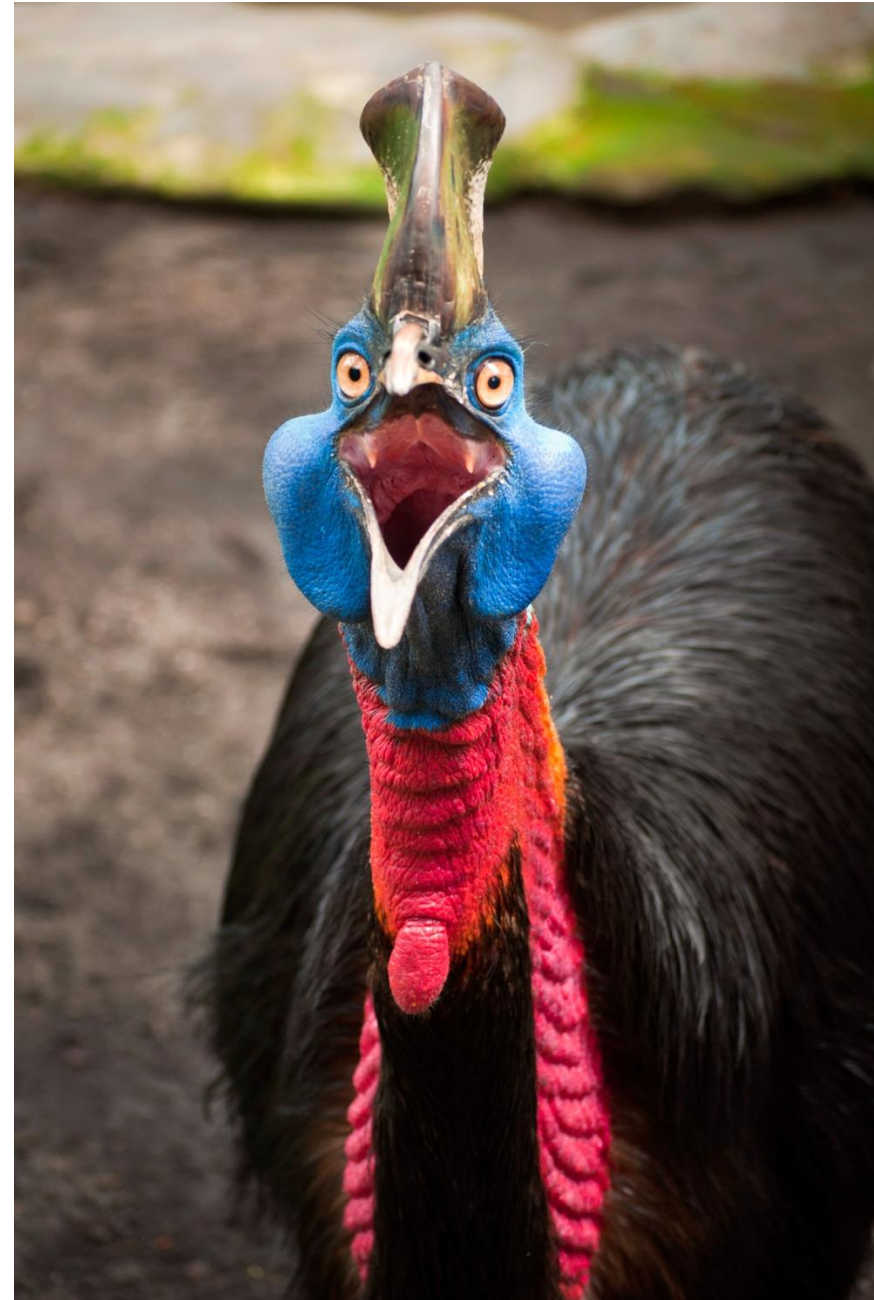
We can also measure attentional compression in the brain

(Mack, Preston, & Love, 2020)



The ability to rapidly reallocate attention should be associated with better learning outcomes. However....

Not all exceptions  
are created equal!



Cassowary (Ghetty Images)



## Crossover exceptions



- Similar to **existing knowledge**, confusable
- Elicit prediction error, cause uncertainty
- Must be distinguished from rule-followers

## Oddball exceptions



- Unique features, distinct from **existing knowledge**
- Novelty without prediction error
- Stored in isolation

The ability to rapidly reallocate attention should be associated with better learning outcomes  
when exceptions overlap with existing knowledge

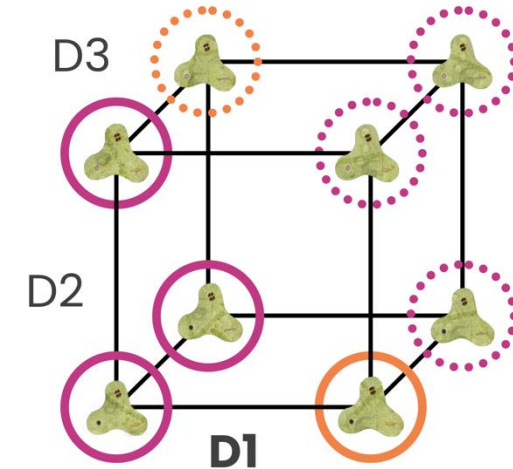
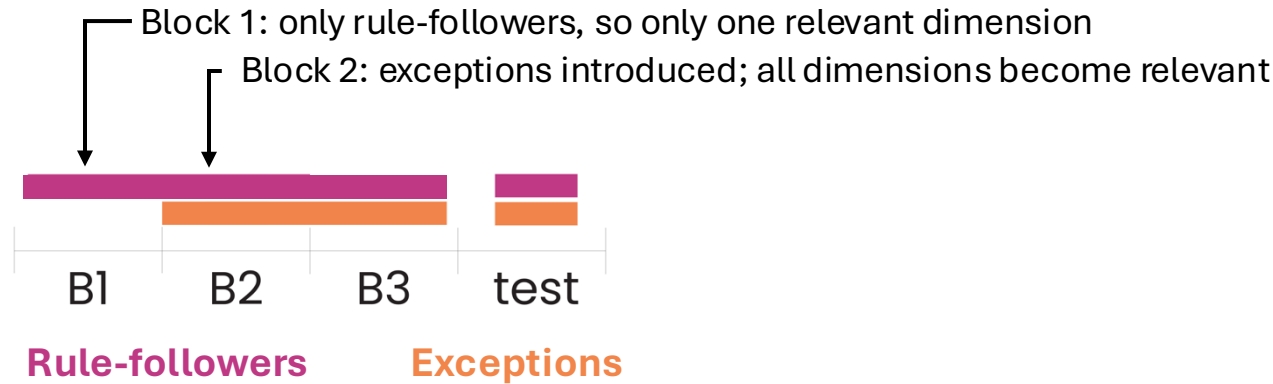
# Experiment design



Xuan (Sophia) Zhang

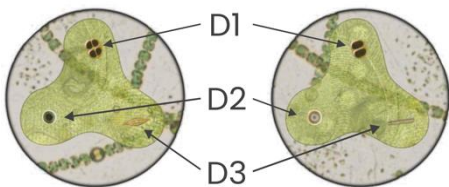


Marian Wang



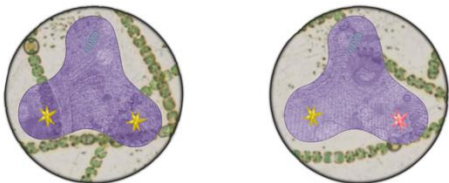
beta gamma

(Shepard, Hovland, & Jenkins, 1961 Type IV)



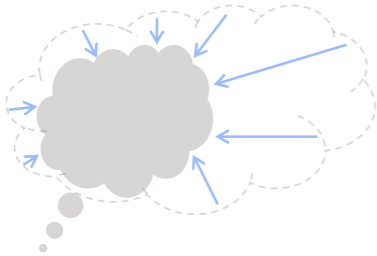
(Blair et al., 2009)

**Crossover exceptions:** Similar to existing knowledge, confusable  
Elicit surprise, prediction error, & uncertainty

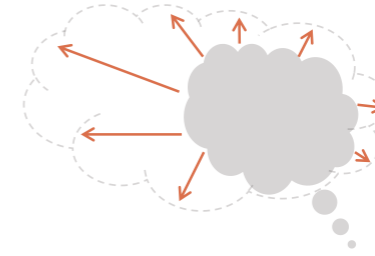
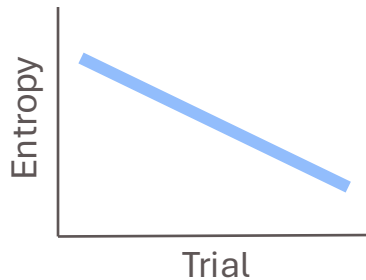


**Oddball exceptions:** Distinct from existing knowledge  
Novelty without prediction error

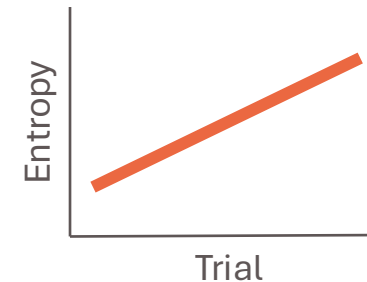
In Information Theory, **entropy (H)** is the amount of information conveyed in an event (Shannon, 1948)



During rule learning, entropy (calculated from mouse-tracking data) should **decrease** as one fixates on fewer features

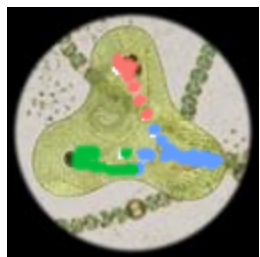


Following (crossover) exception introduction, entropy should **increase** as attention expands

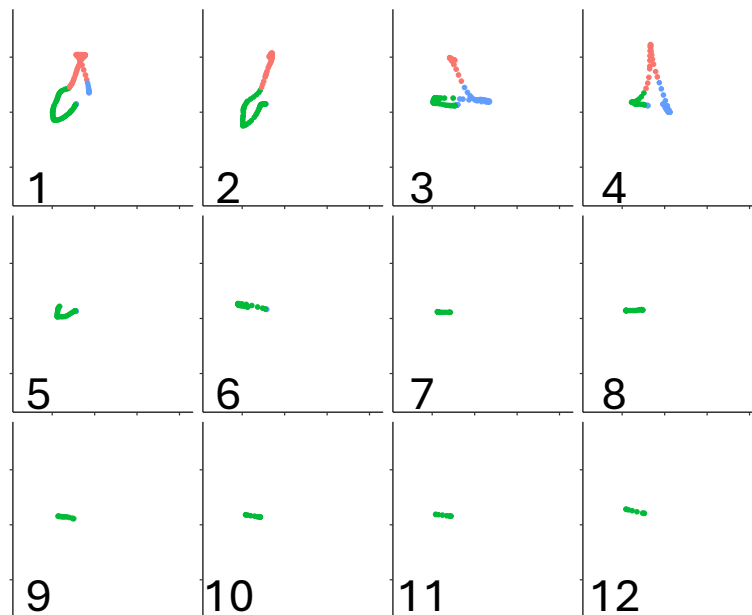


**We can use entropy to track attentional tuning before and after exception introduction**





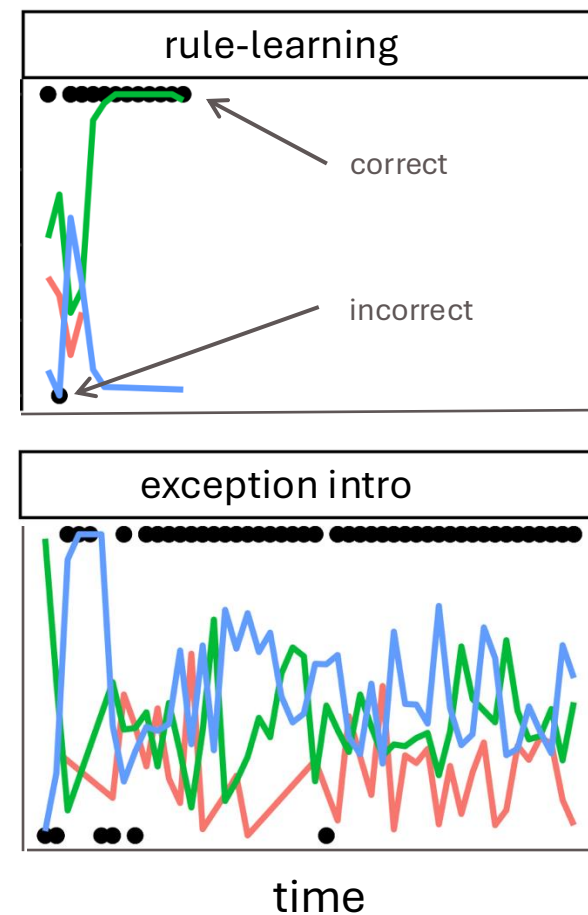
Mouse-tracking  
data



Trial-by-trial  
(rule learning)



fixation proportion



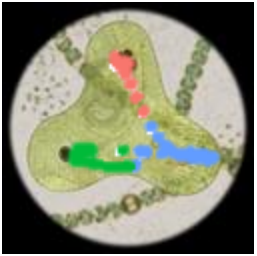
High entropy



Low entropy

# Only confusable exceptions elicit widespread shifts in attention

## Entropy as a measure of attentional distribution



High entropy    Low entropy

## Crossover introduction

LB1: Attention contraction

LB2: Widespread attention expansion

LB3: Attention optimization

## Oddball introduction

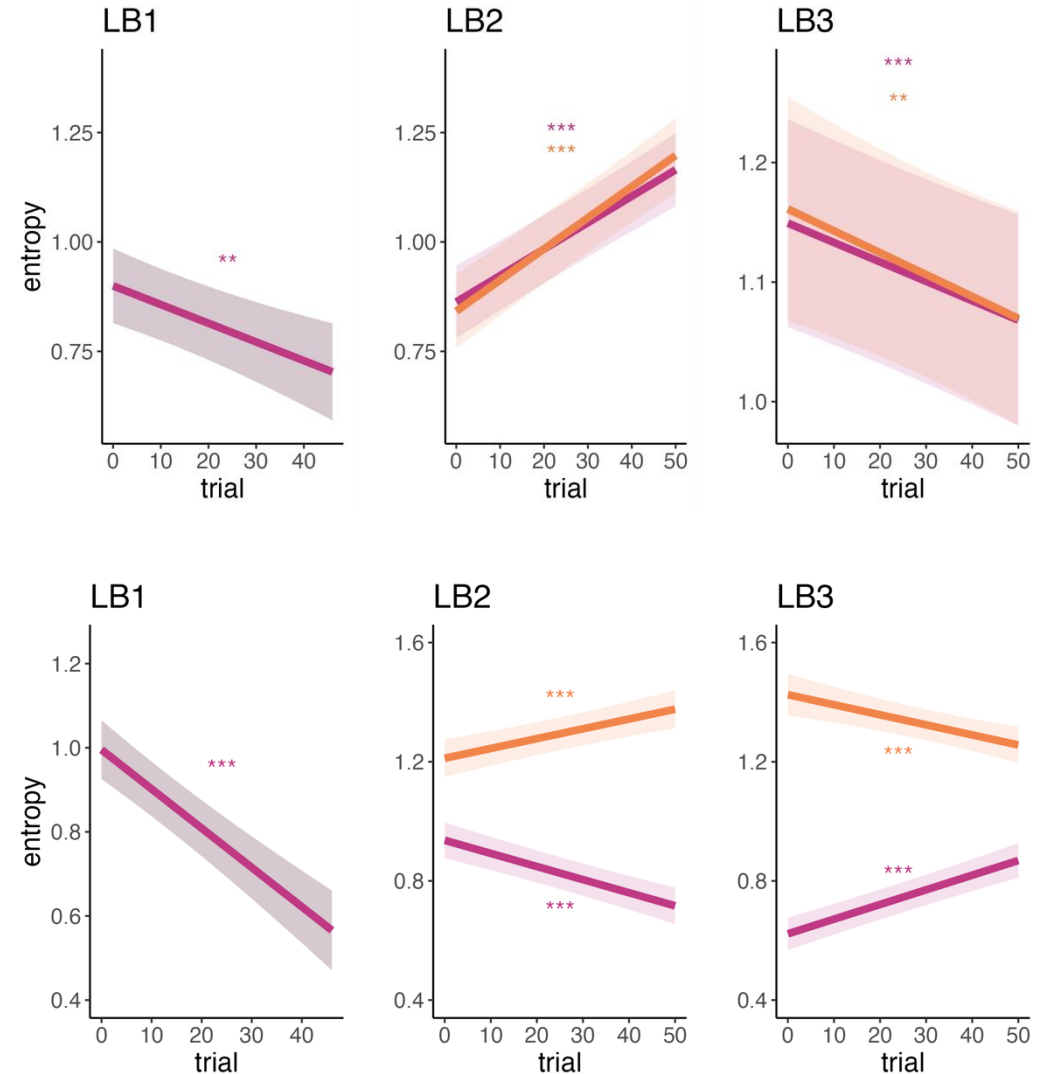
LB1: Attention contraction

LB2: Exception-specific expansion

LB3: Potential fatigue?

rule-followers

exceptions

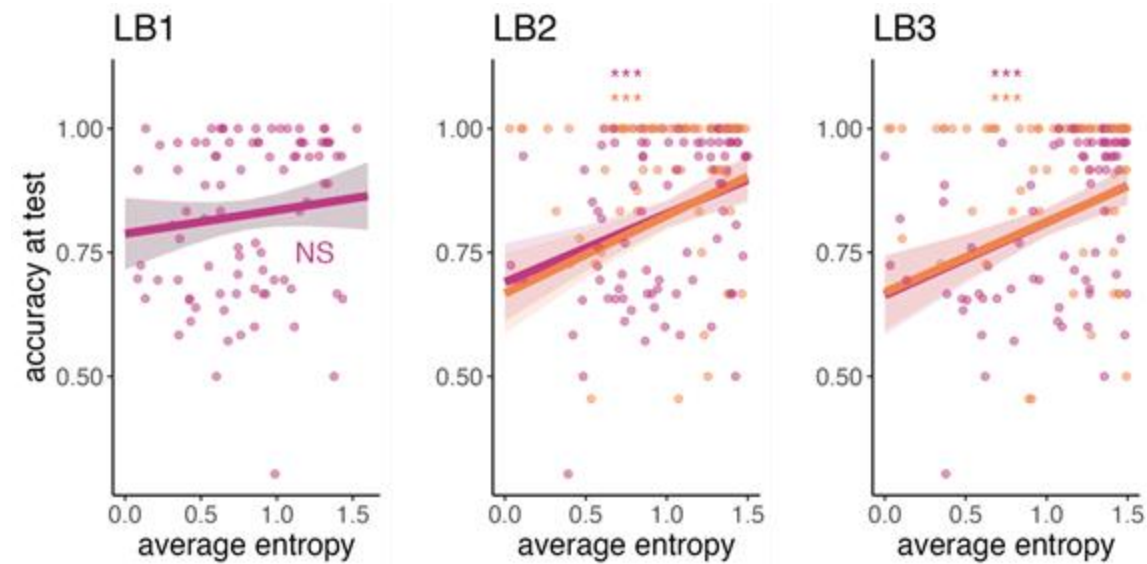


# Learning confusable exceptions demands attention expansion

## Crossover introduction

LB1: No relation

LB2&3: Increased entropy  
associated with  
increased accuracy

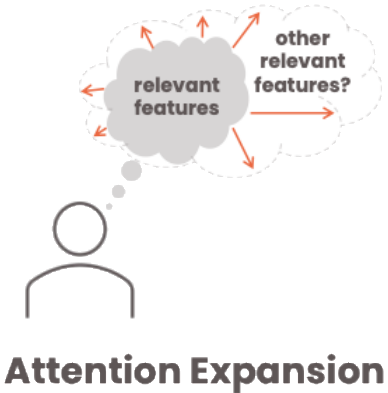
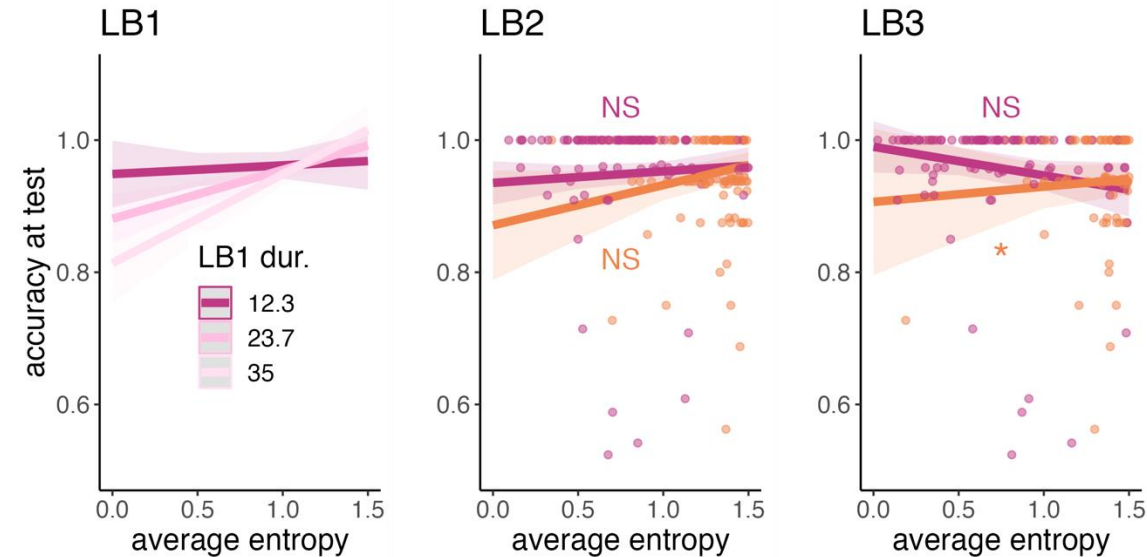


## Oddball introduction

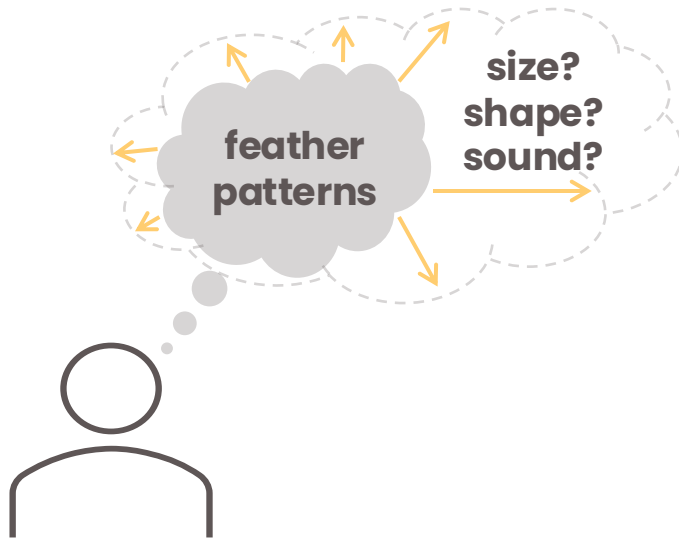
LB1: Duration-dependent  
relation

LB2: No relation

LB3: Positive relation for  
exceptions



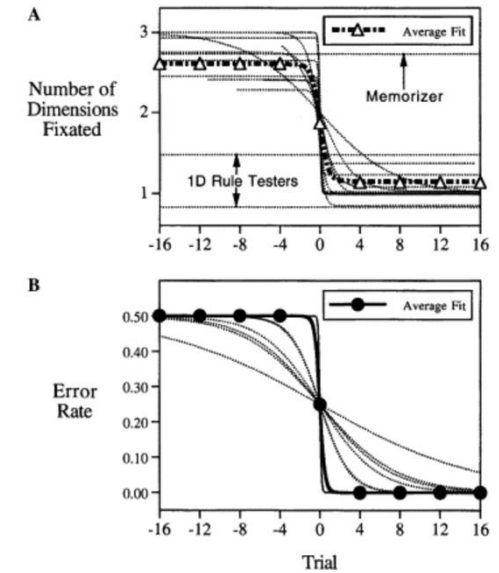
# Some Takeaways



Crossover (but not oddball) exceptions lead to adaptive expansion of attention



How surprising information overlaps with existing knowledge is important for attention and learning



This work corroborates existing findings on how uncertainty mediates selective attention.

# Thank you!



<http://macklab.utoronto.ca>

Michael Mack	Yongzhen (Dory) Xie
Frida Printzlau	Sanjivan Lognathan
Sagana Vijayarajah	Selina Fu
Gaeun Son	Rafidal Islam
Mateja Perovic	Marian Wang
Melisa Gumus	Xuan (Sophia) Zhang

And members of Budding Minds Lab

