85426/85726: Learning in Humans and Machines

Professor Dan Yurovsky

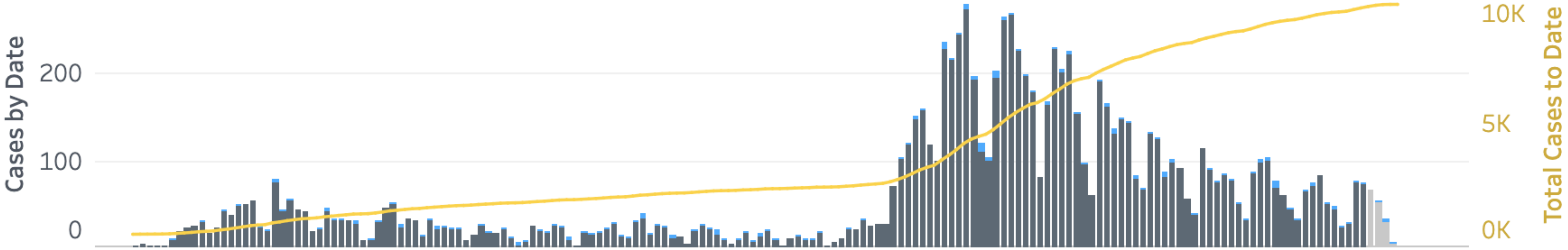
Fall 2020

Welcome to Fall 2020!

Things are definitely a little scary right now…

https://www.cmu.edu/piper/news/archives/2020/august/campus-will-be-different.html

Try to find optimism where you can



https://www.alleghenycounty.us/Health-Department/resources/COVID-19/Covid-19.aspx

Where I’m from

https://www.cmu.edu/dietrich/psychology/callab/papers/ysy-devsci-2013.pdf

I did my undergrad in Computer Science, Cognitive Science, and Human-Computer Interaction at CMU

I got my PhD in Cognitive Psychology at Indiana University studying “statistical learning” as a model of children’s language learning

Spent some time at Stanford and the University of Chicago, and now here I am

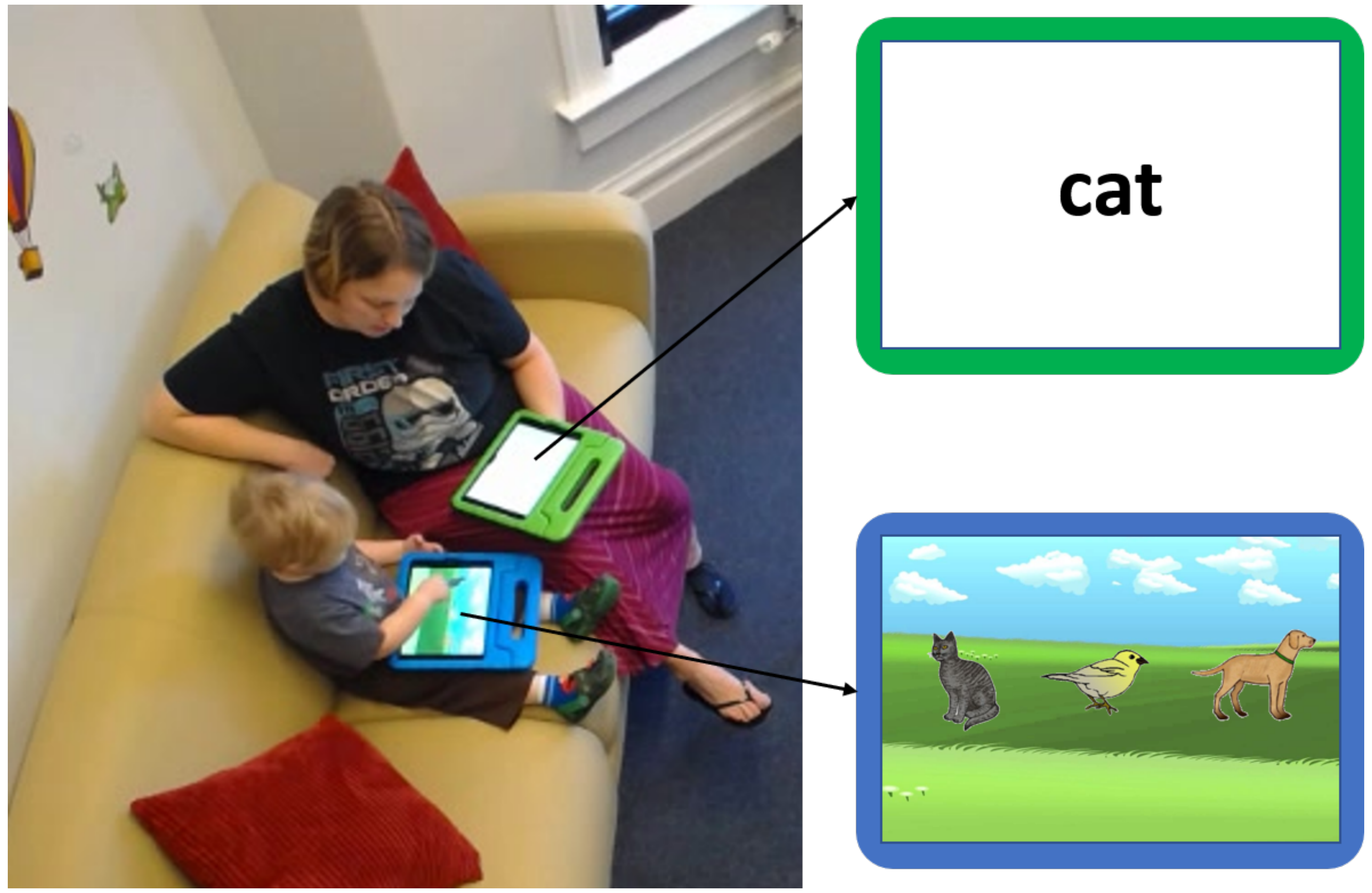
What I do

I'm a faculty member in the Psychology department.

Folks in my lab study how people learn—especially from other people.

How do the interactions between parents and their children give rise to language learning?

Communication and Learning Lab



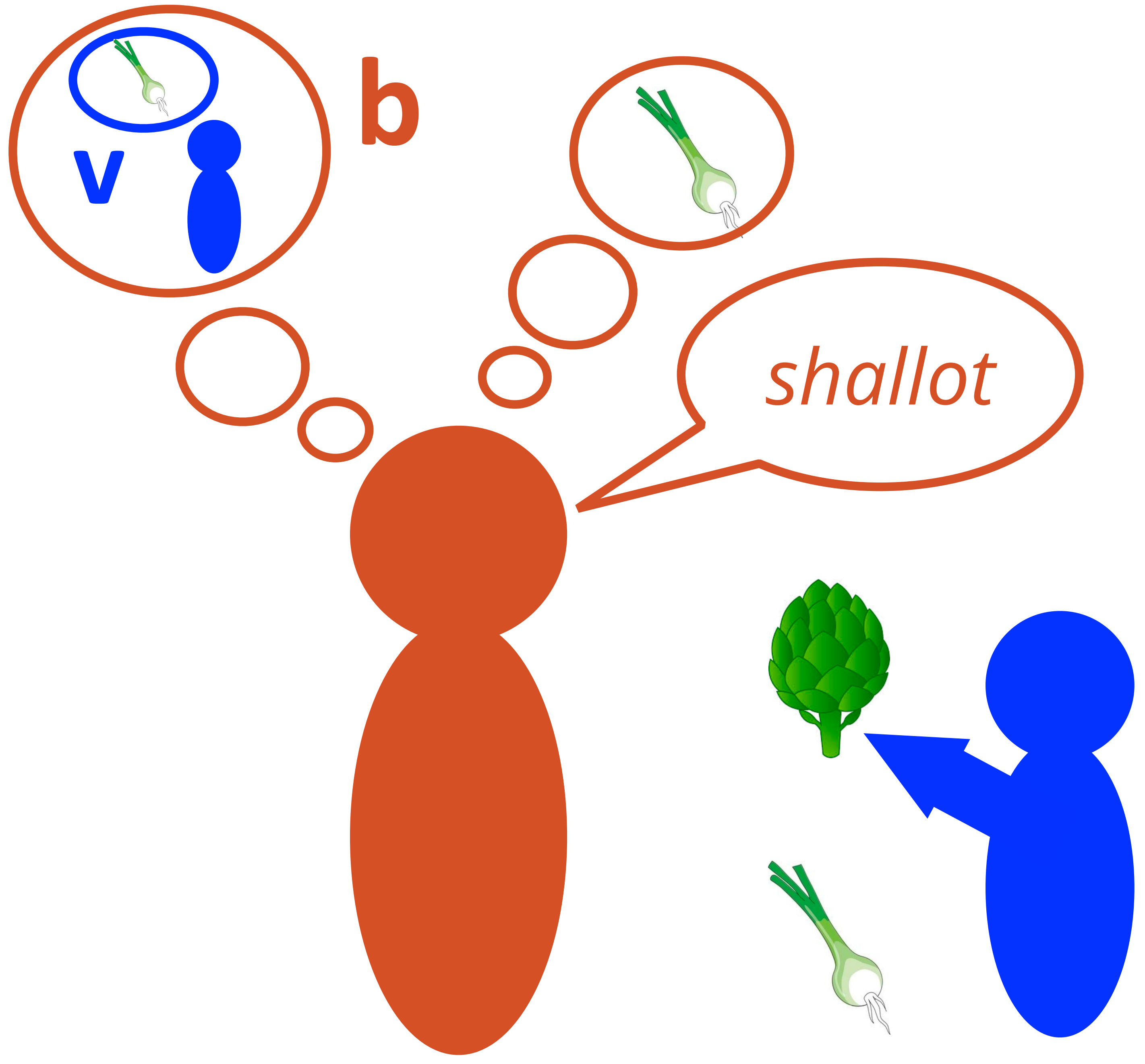
https://www.cmu.edu/dietrich/psychology/callab/papers/leung_cogsci2019.pdf

Why I'm here

I use computational models to understand how people learn.
I think you should too!

Experiments and formal models are the two best tools we have for making sense of how people behave.

And sometimes models teach us about how to build machines that solve the same problems people do.



https://www.cmu.edu/dietrich/psychology/callab/papers/morris_cogsci2019.pdf

Who are you? Why are you here?

A look at the schedule and syllabus

Office hours poll

How does learning work? Let's make an affinity diagram

http://bit.ly/learning-affinity

Three awesome things that models can do

1. **Models can clarify our assumptions (which often turn out to be wrong)**

2. **Models are an efficient way to compressing a lot of experimental data**

3. **Models can help figure out when two different accounts make identical predictions**

Three awesome things that models can do

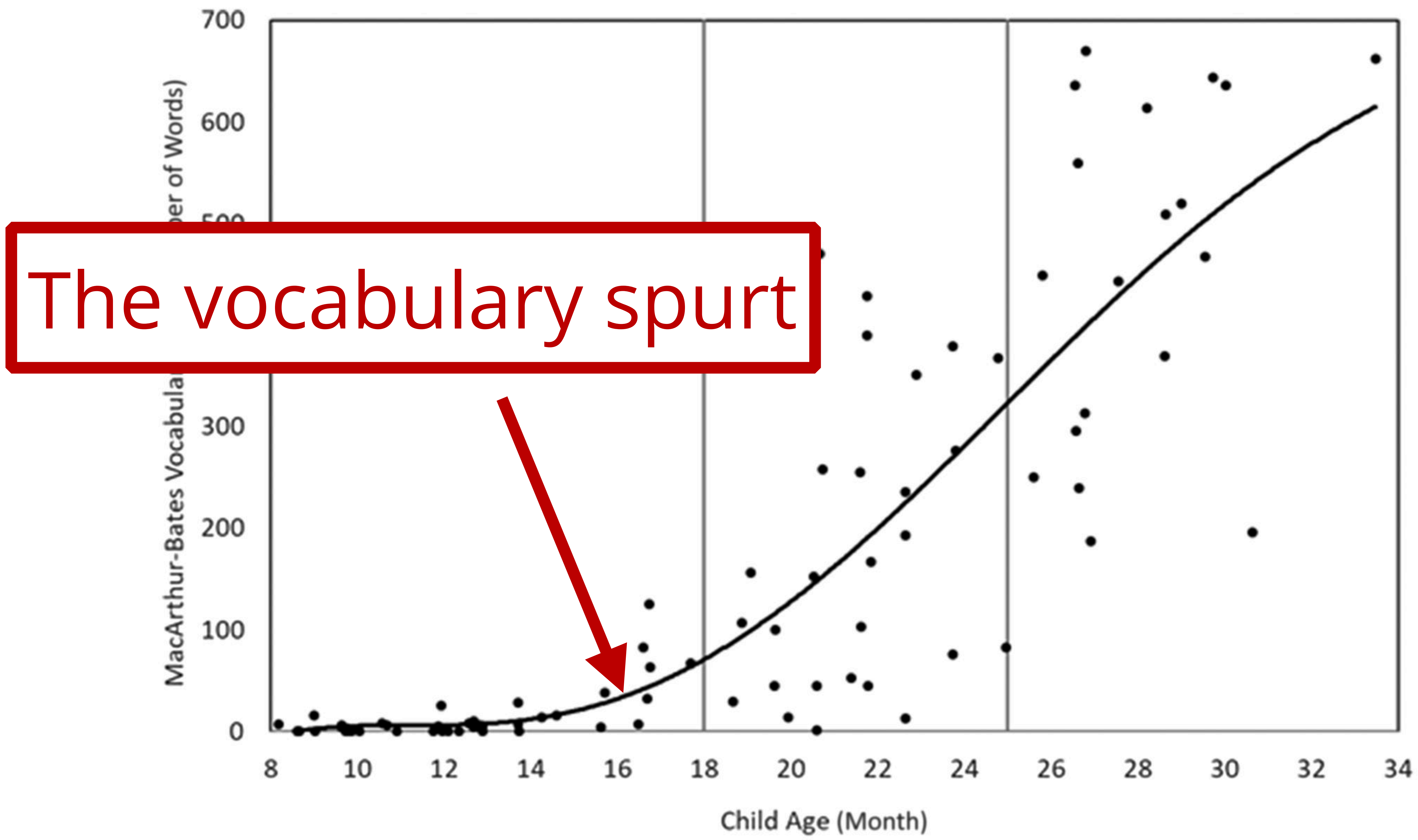
1. **Models can clarify our assumptions (which often turn out to be wrong)**

2. Models are an efficient way to compressing a lot of experimental data

3. Models can help figure out when two different accounts make identical predictions

The childhood vocabulary spurt

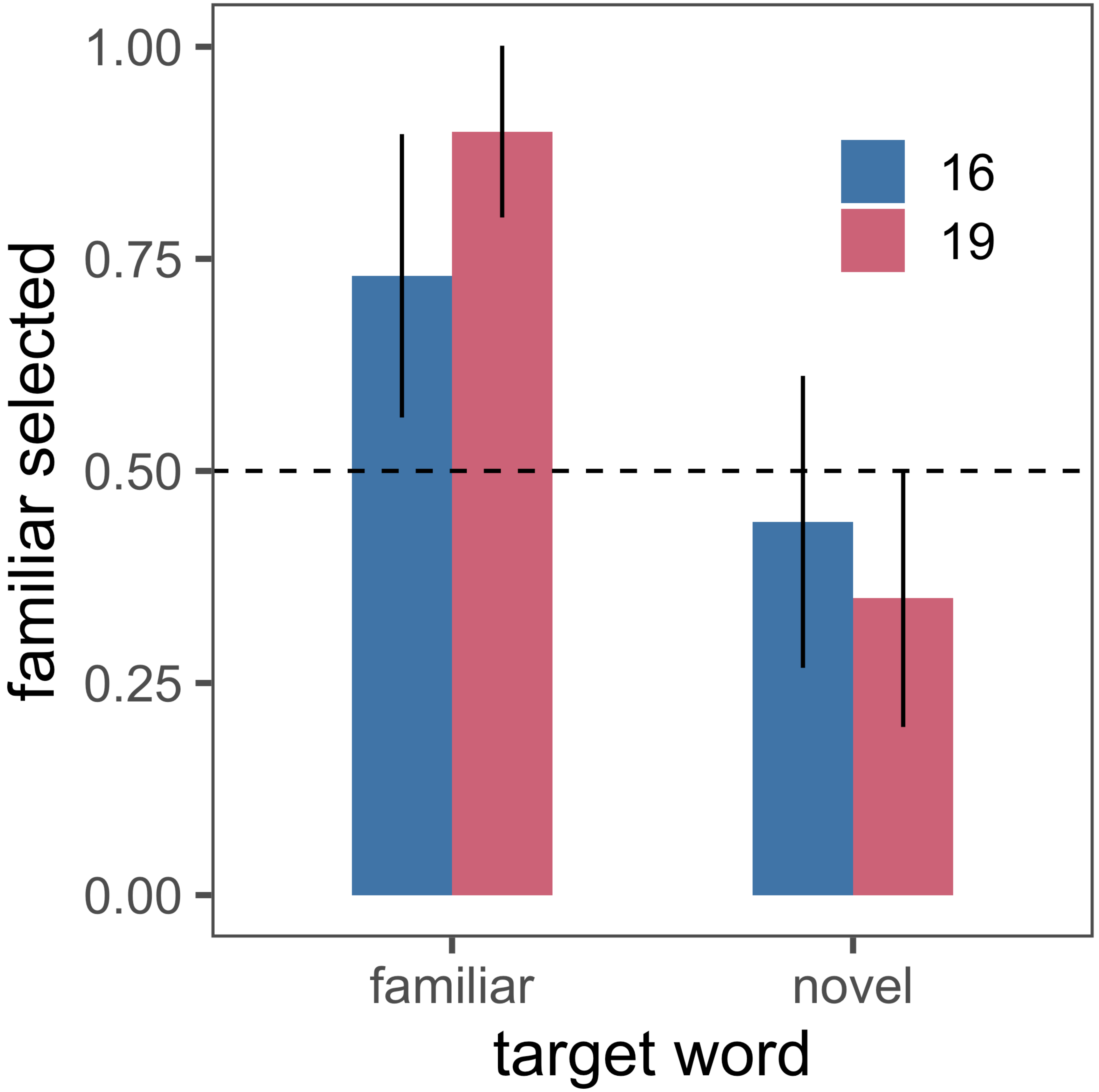
Sometime around 16 months, children start to learn words at a much faster rate



Seems like something about these children must have changed!

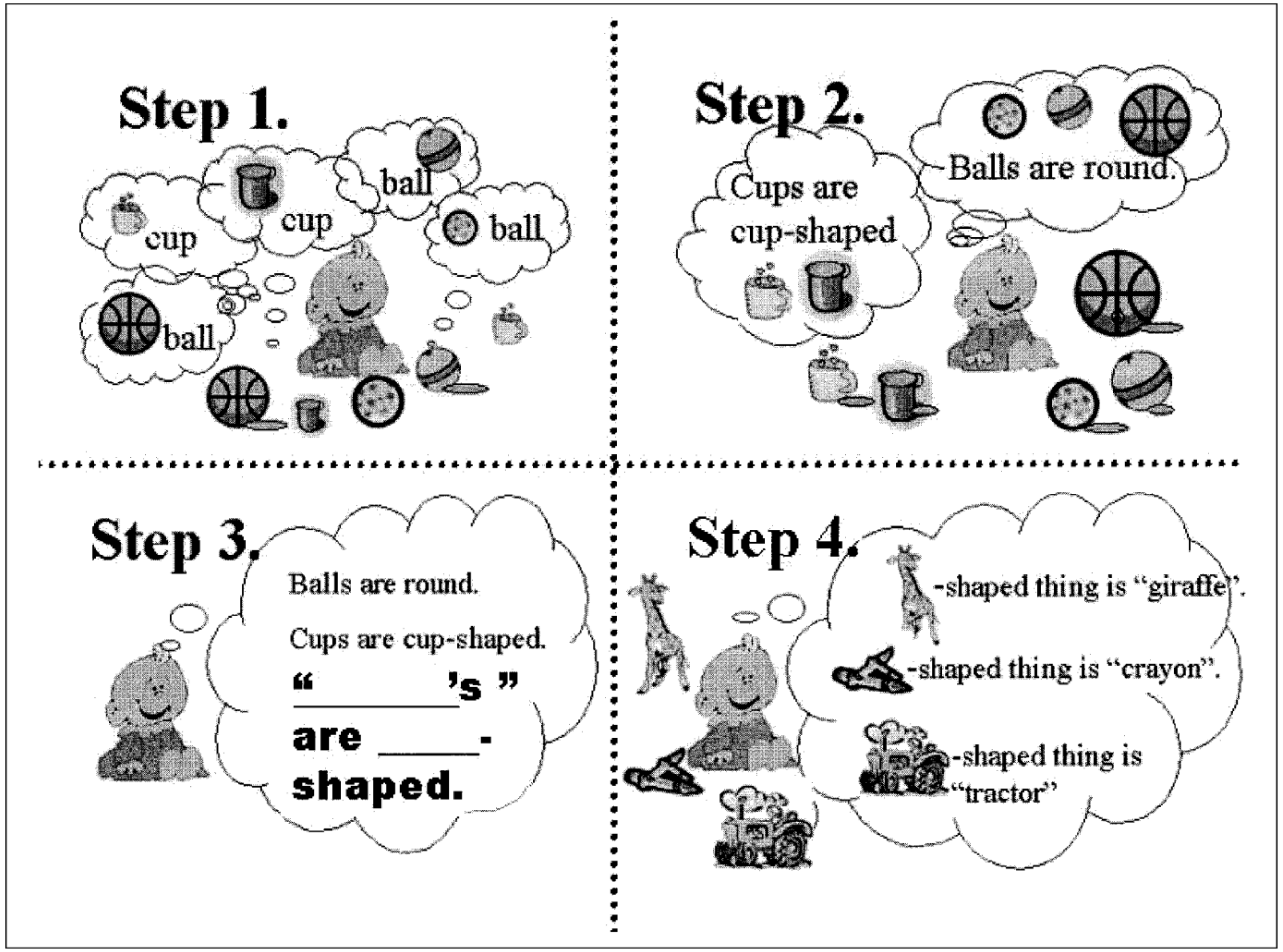
Any idea what it might be?

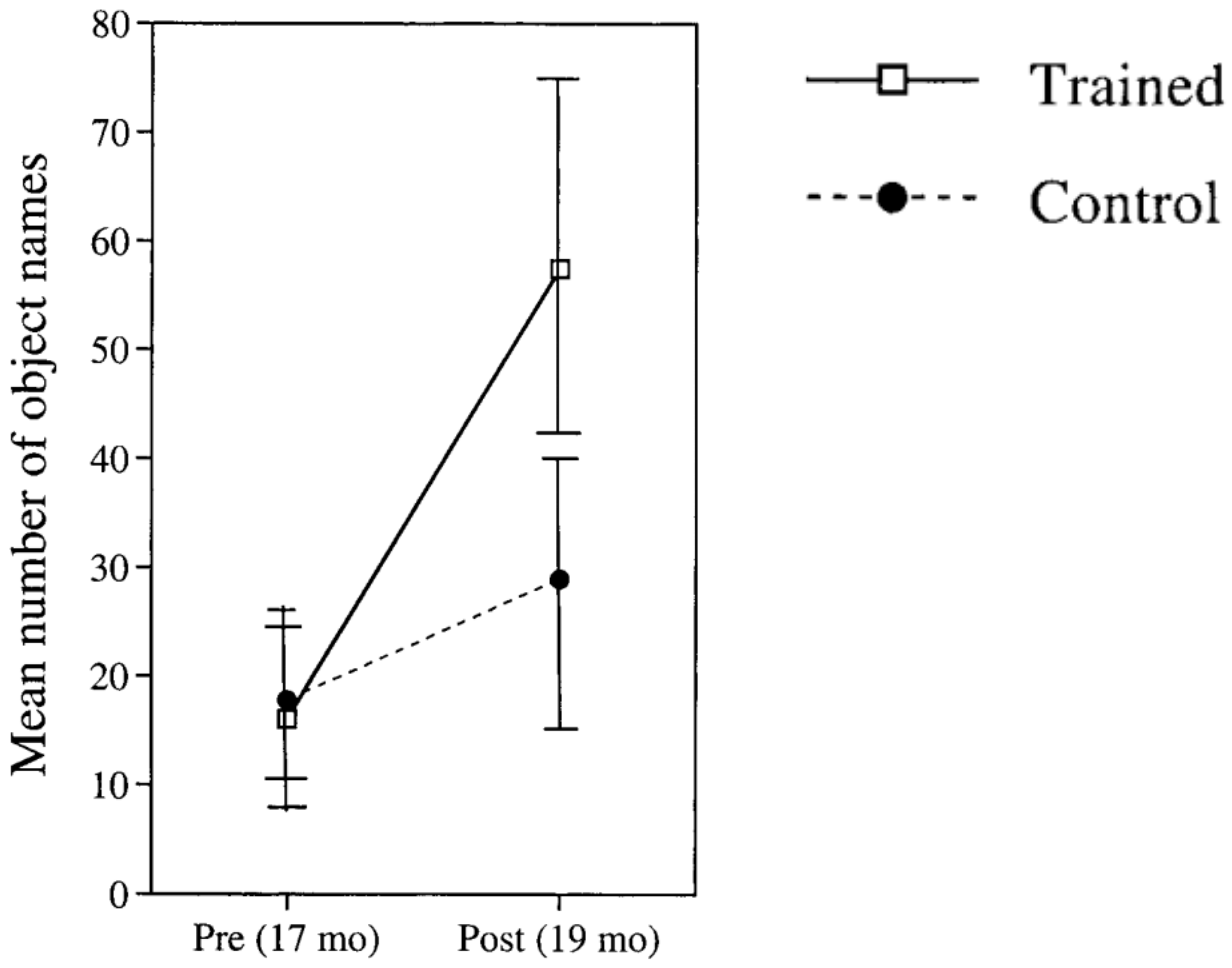
One hypothesis: Mutual exclusivity - Markman et al. (2003)



Can you find the **dax**?

Another hypothesis: Shape bias - Smith et al. (2002)





One plausible answer: Nothing at all!

Because we saw a change in behavior, we assumed that there must be a change in the underlying system

But could this “spurt” is the Null Hypothesis?

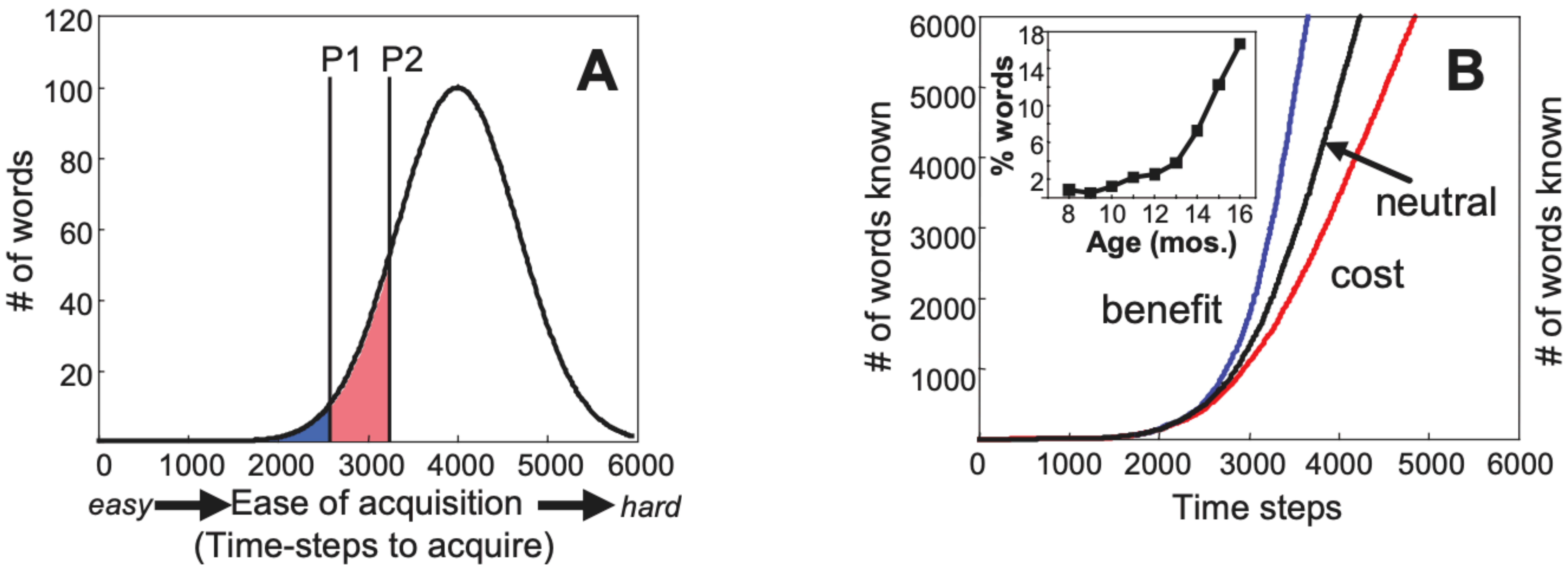
McMurray (2007). Defusing the childhood vocabulary explosion.

https://science.sciencemag.org/content/sci/317/5838/631.full.pdf

Accumulator models of learning

Implementing McMurray’s basic model

The vocabulary spurt as a null hypothesis



What does this model tell us?

Three awesome things that models can do

1. Models can clarify our assumptions (which often turn out to be wrong)

2. **Models are an efficient way to compressing a lot of experimental data**

3. Models can help figure out when two different accounts make identical predictions

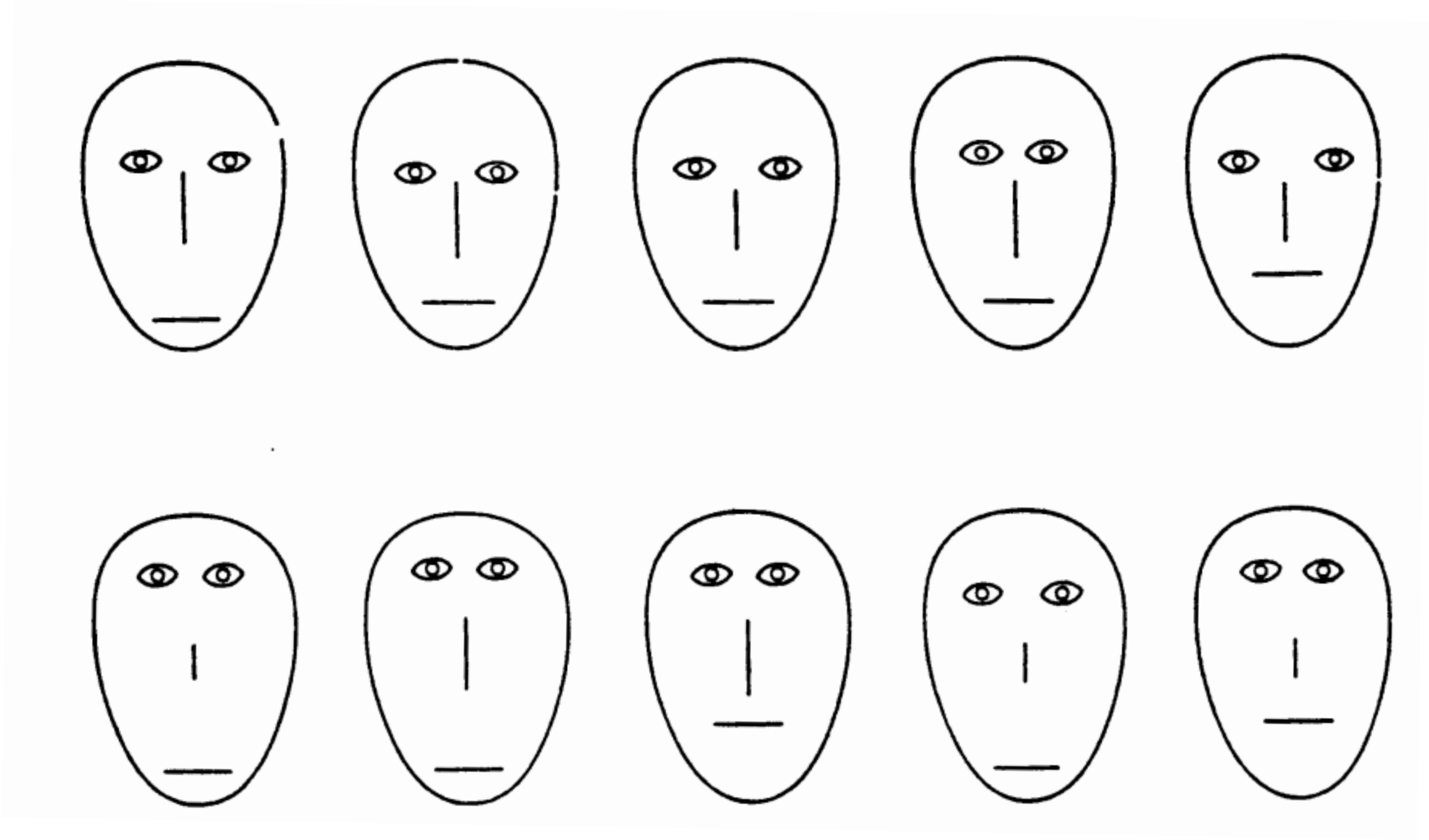
The rational model of categorization - Anderson (1991)

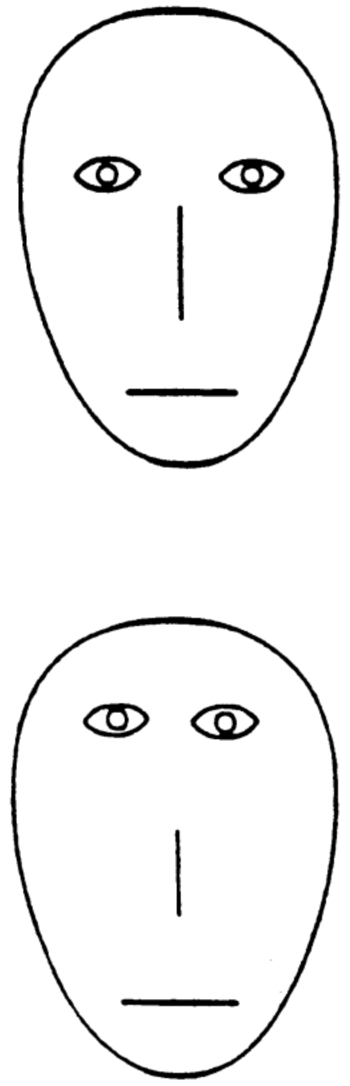
The Adaptive Nature of Human Categorization

John R. Anderson
Carnegie Mellon University

A rational model of human categorization behavior is presented that assumes that categorization reflects the derivation of optimal estimates of the probability of unseen features of objects. A Bayesian analysis is performed of what optimal estimations would be if categories formed a disjoint partitioning of the object space and if features were independently displayed within a category. This Bayesian analysis is placed within an incremental categorization algorithm. The resulting rational model accounts for effects of central tendency of categories, effects of specific instances, learning of linearly nonseparable categories, effects of category labels, extraction of basic level categories, base-rate effects, probability matching in categorization, and trial-by-trial learning functions. Although the rational model considers just 1 level of categorization, it is shown how predictions can be enhanced by considering higher and lower levels. Considering prediction at the lower, individual level allows integration of this rational analysis of categorization with the earlier rational analysis of memory (Anderson & Milson, 1989).

The rational model of categorization - Anderson (1991)

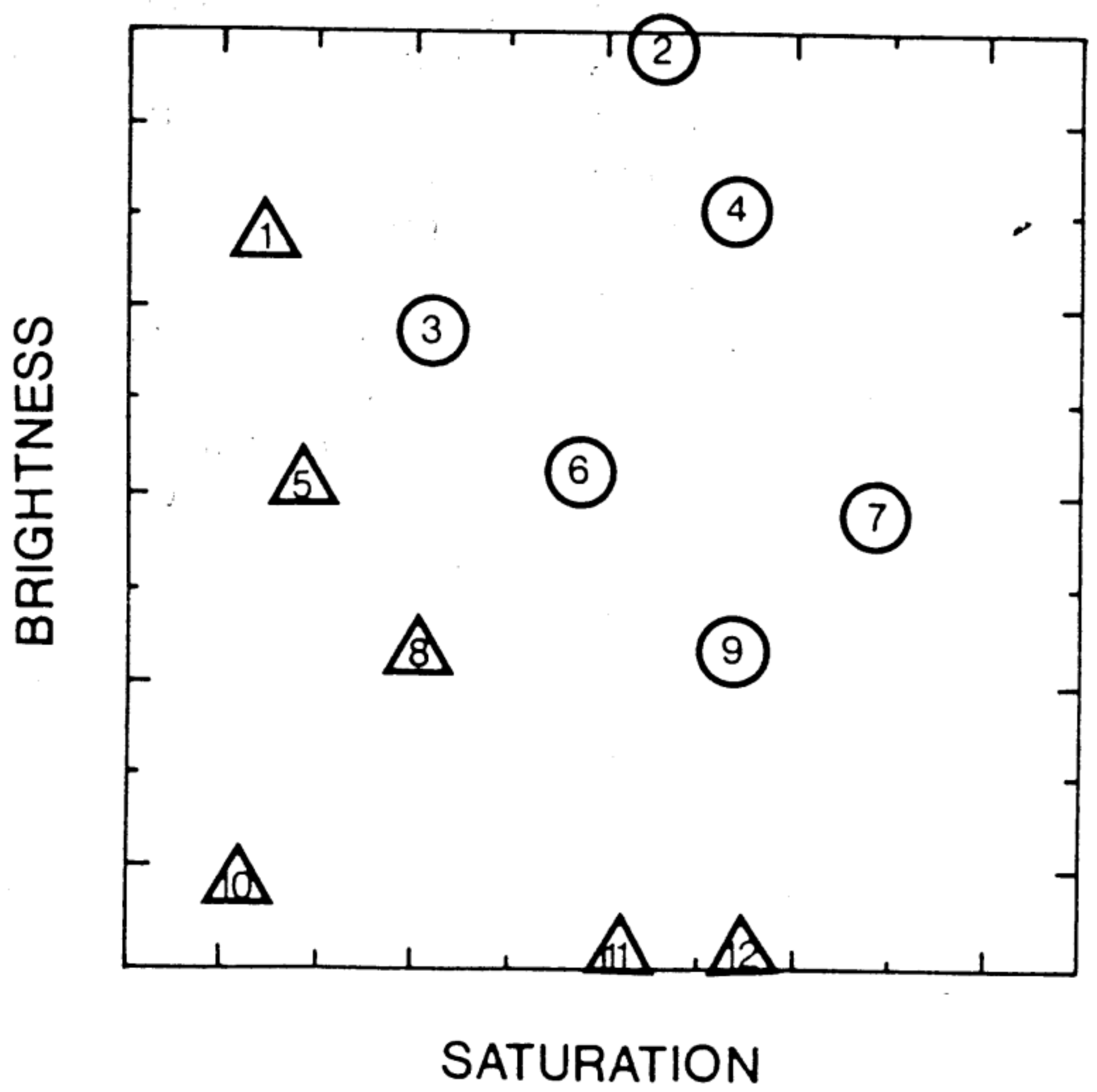


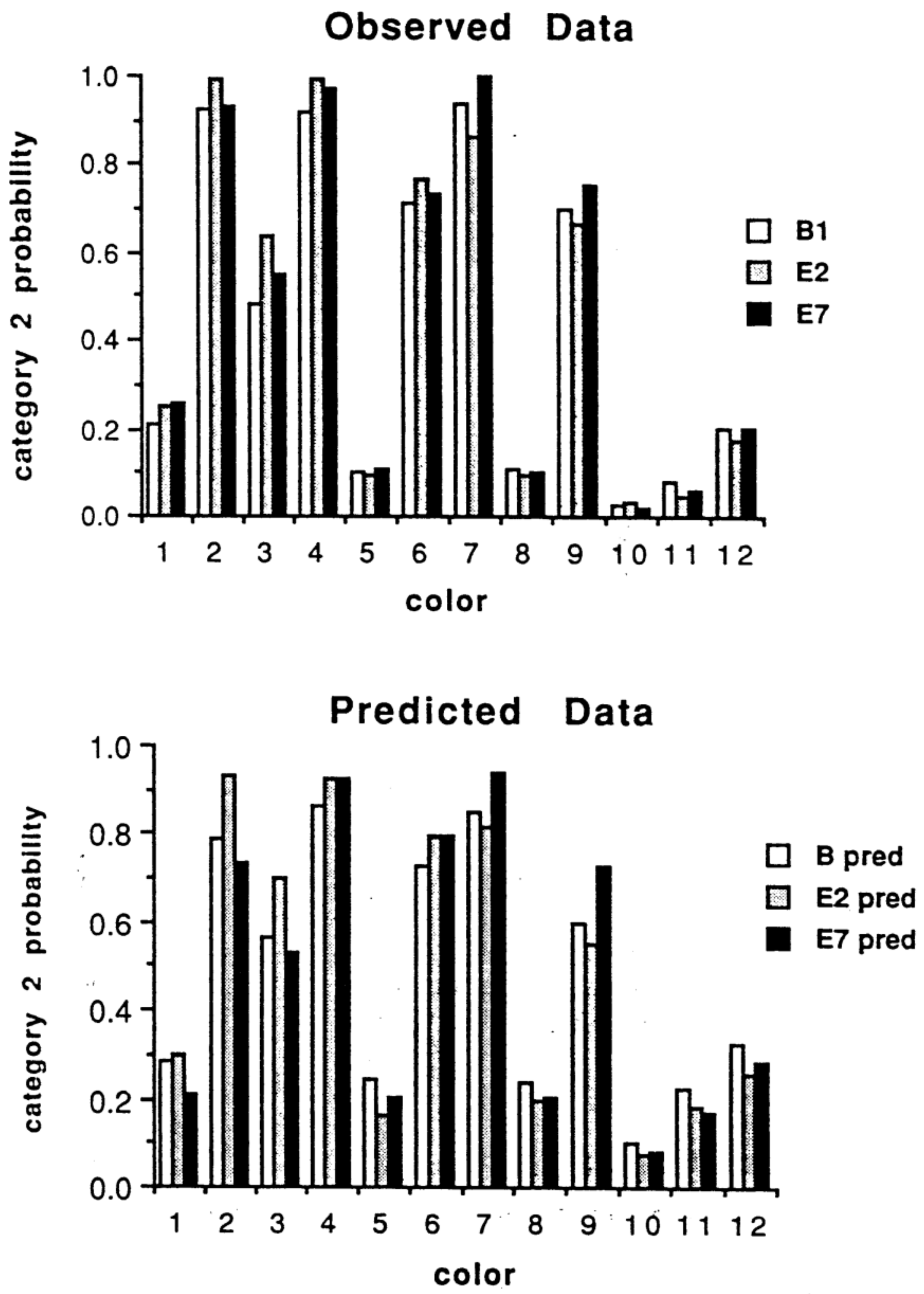


Categorization accuracy varies with distance from the prototype

Reed (1972)

The rational model of categorization - Anderson (1991)

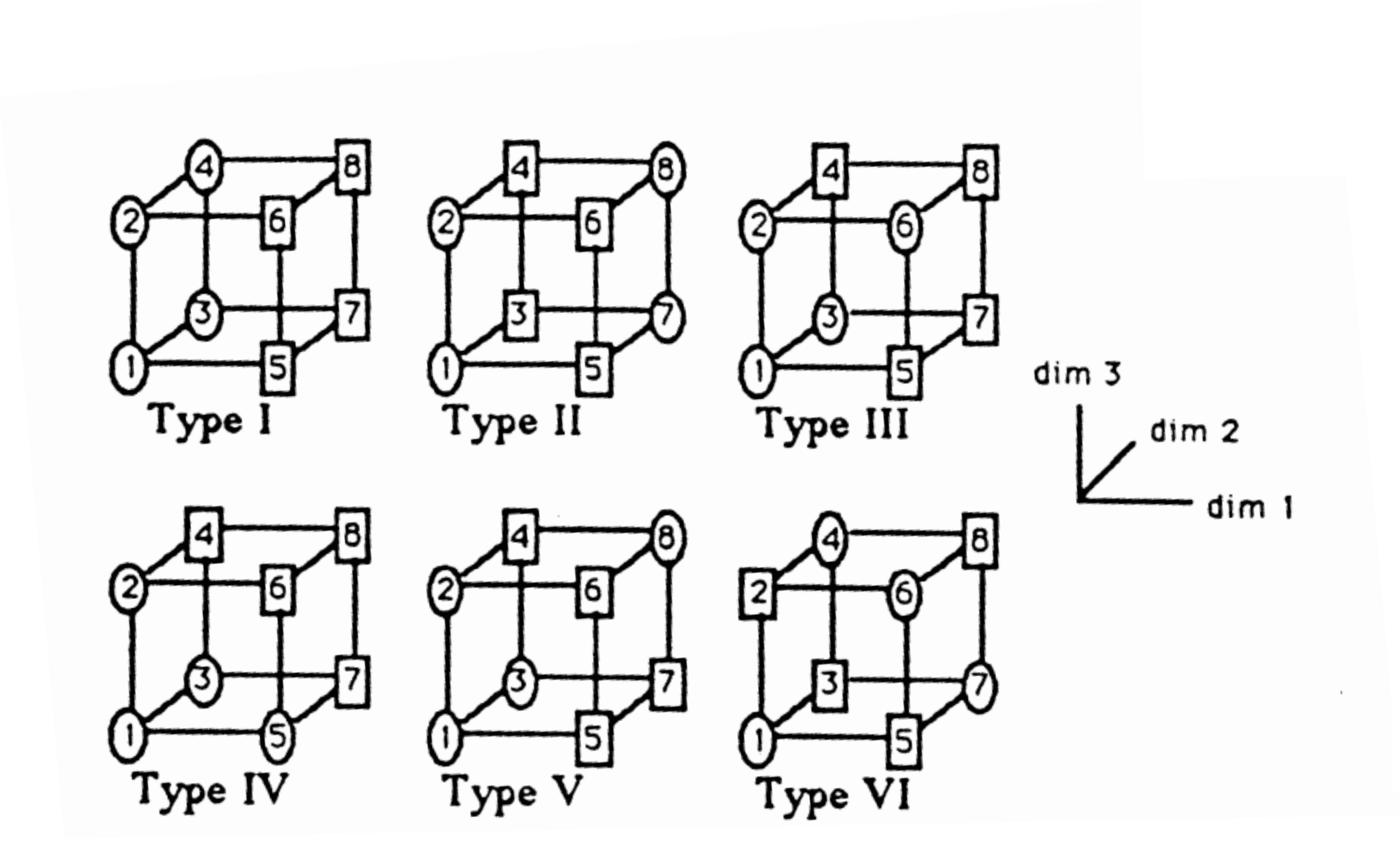


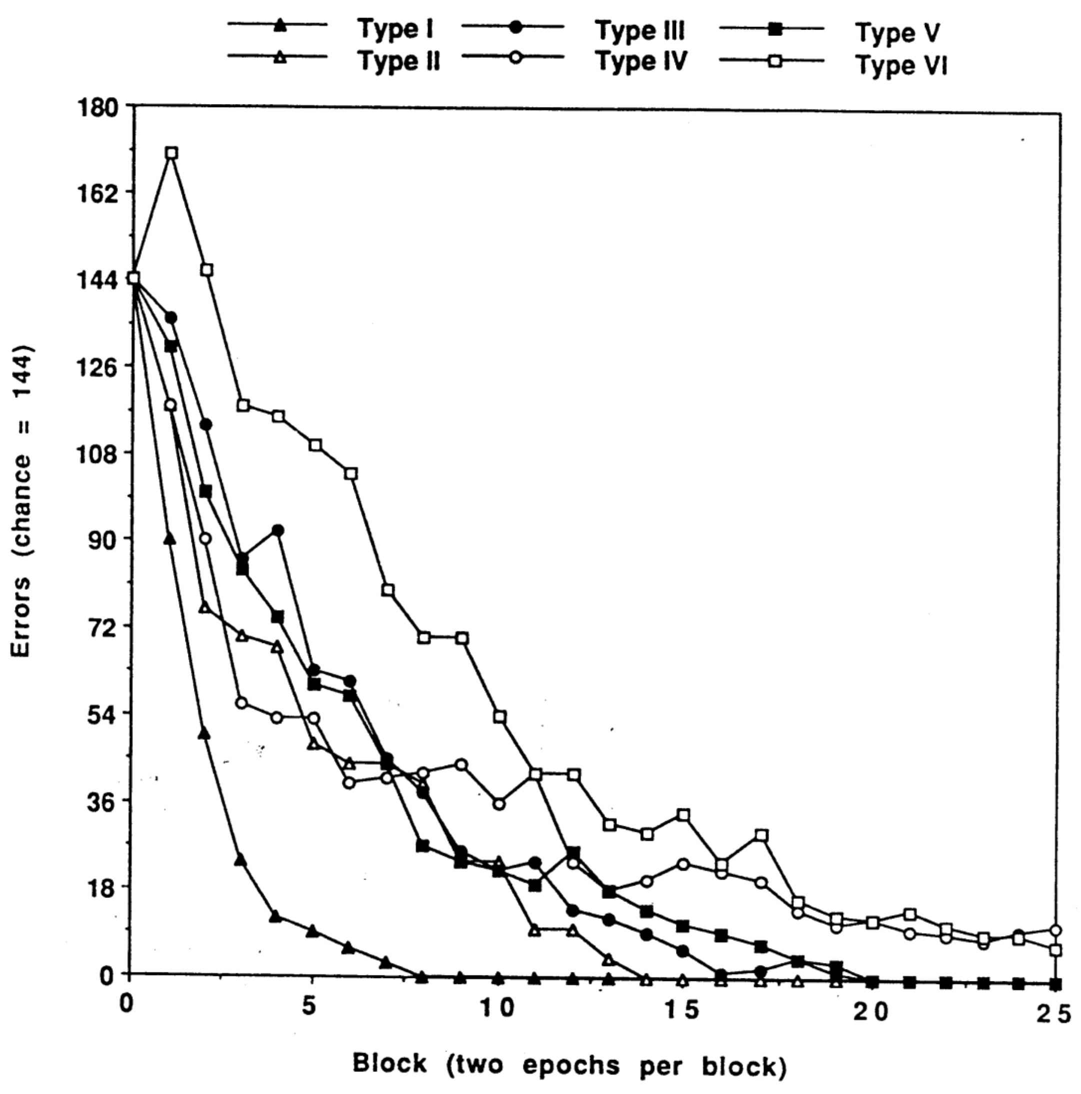


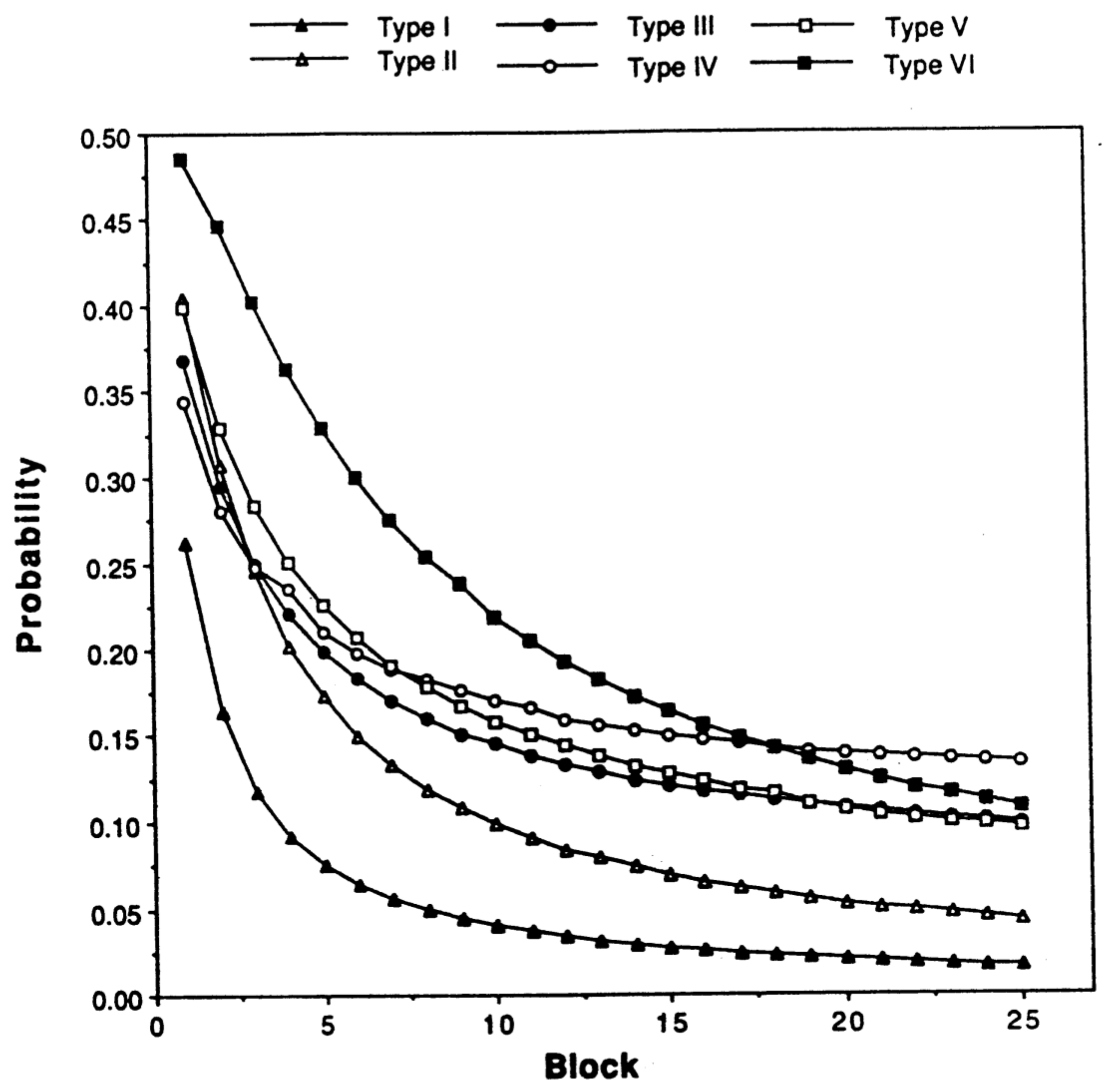
Exemplars closer to category boundaries are harder to learn

Nosofsky (1988)

The rational model of categorization - Anderson (1991)







Complex rules are harder to learn than simple rules

Shepard et al. (1961)

Side benefits of models as compression algorithms

1. **Trying to compress a lot of data can tell us when models are *inconsistent***
   - Parameters of the same mechanism (e.g. learning rate, attention, etc.) sometimes fit different experiments at *different* levels, but not all of the experiments at the *same* level

2. **Model parameters let us transform manipulations of experimental variables into manipulations of cognitive variables**

Three awesome things that models can do

1. Models can clarify our assumptions (which often turn out to be wrong)

2. Models are an efficient way to compressing a lot of experimental data

3. **Models can help figure out when two different accounts make identical predictions**

Very different can be indistinguishable - Townsend (1972)

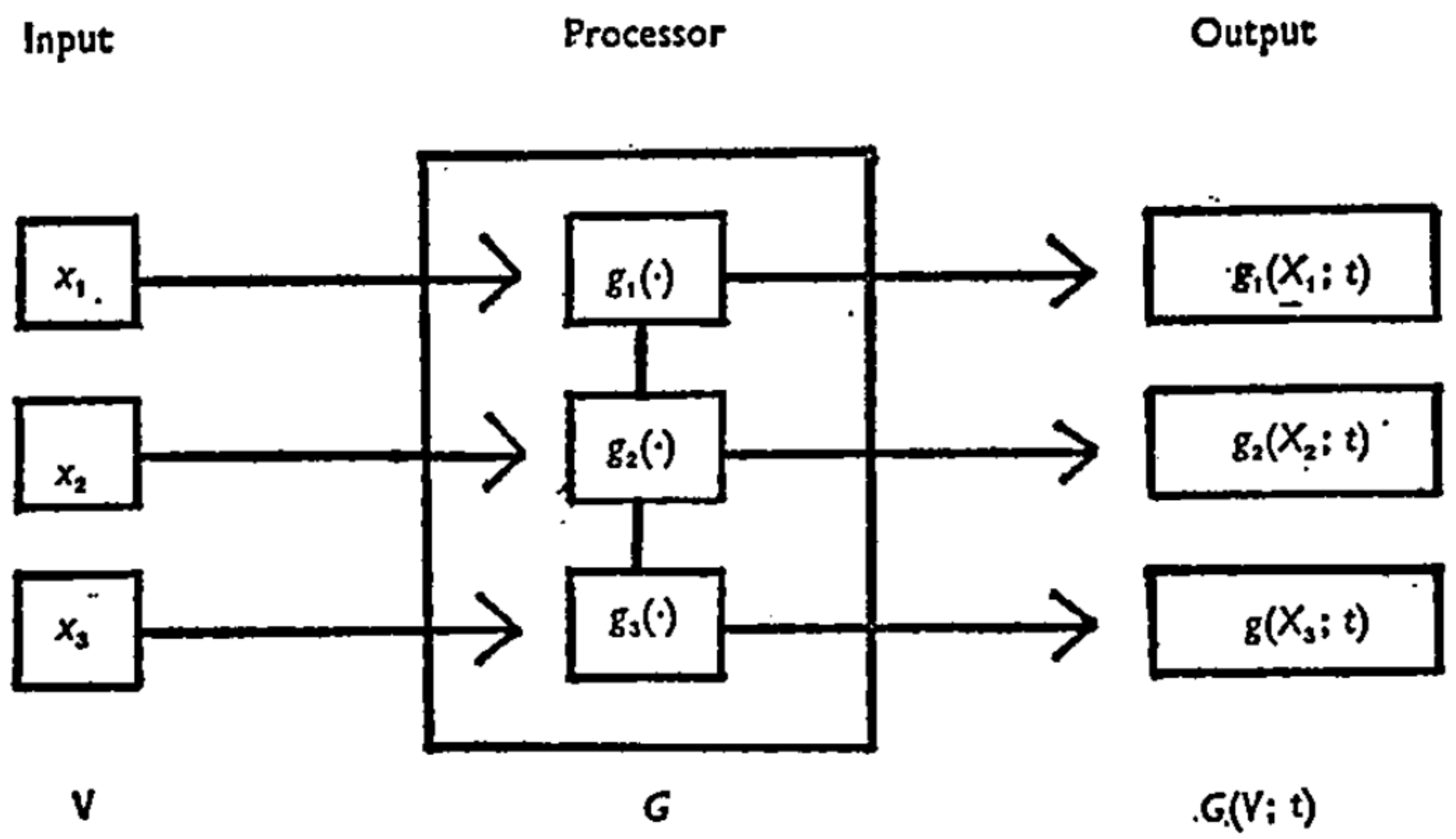
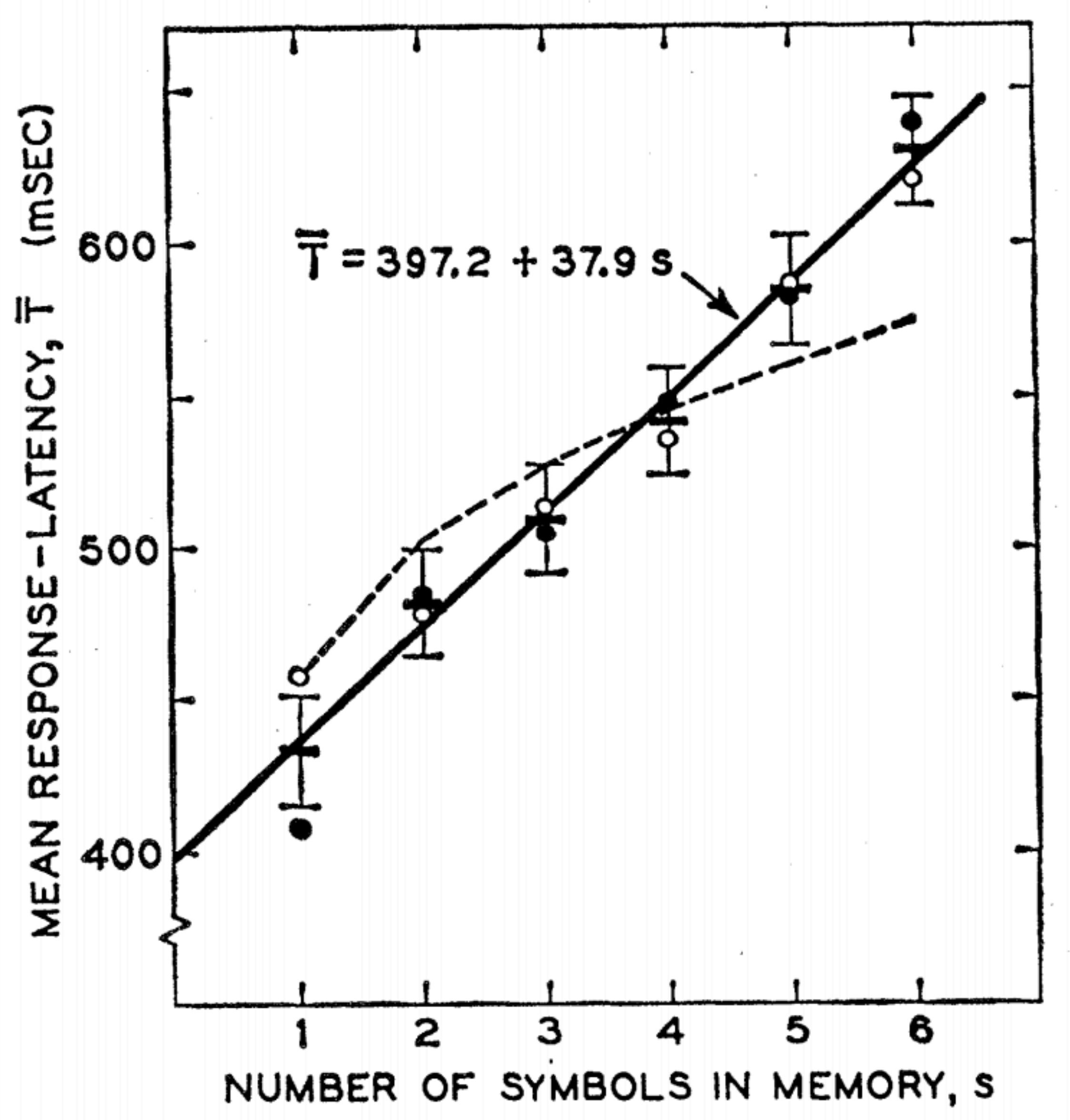


FIG. 1.—Schematic for a parallel processor.

A C V F G



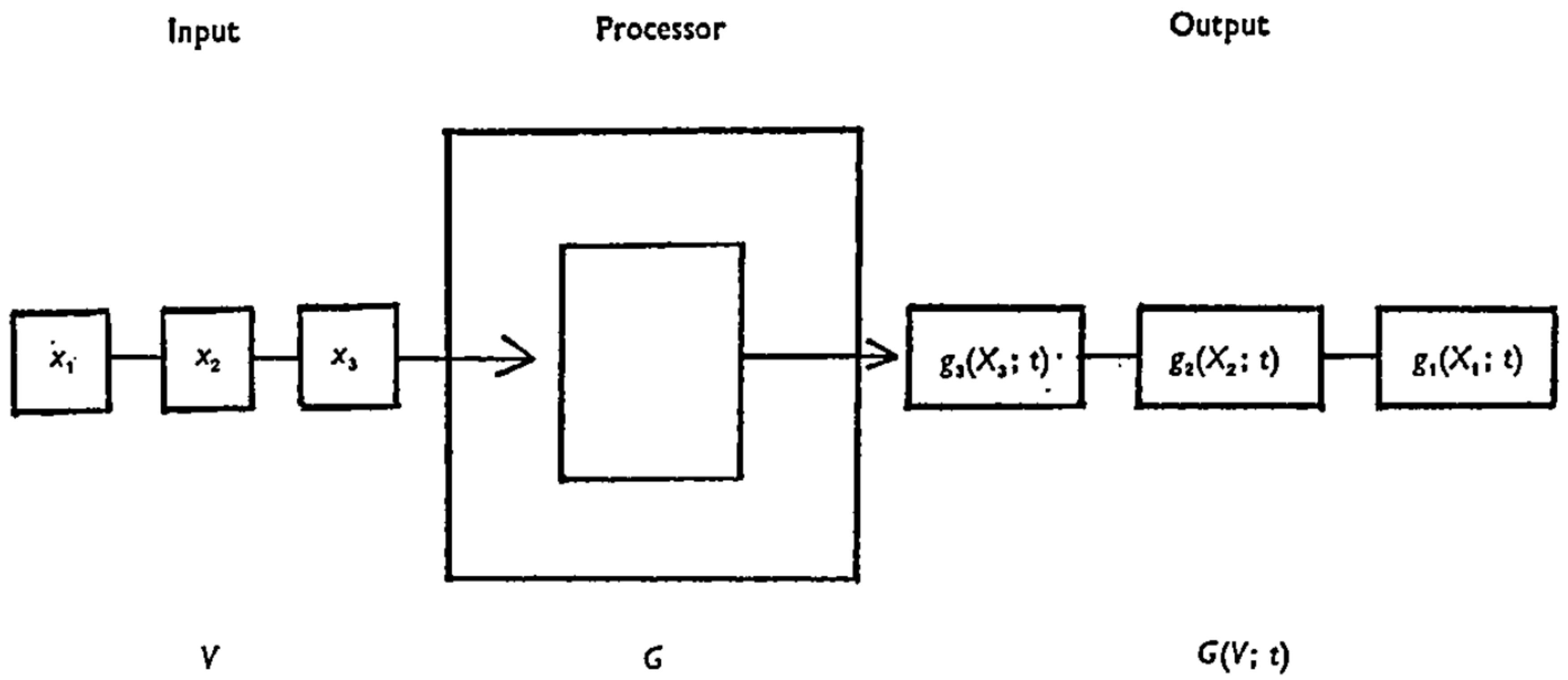


FIG. 2.—Schematic for a serial processor.

Looking forward two weeks

	Simple Neural Networks	Reading	Assignment	Class
September 1	**Learning in Humans and Machines**			
September 3	**R, RStudio, and Github**			
September 8	**Associative Learning**			
September 10	**How far can simple associative learning get you?**			
September 15	**Indirect associations**			
September 16	Implementing the Rescorla-Wagner Model			

bit.ly/lhm-feedback