Title: Theory formation, causal models, and the evolution of learning

Date: Sep 27, 2017 02:00 PM

URL: http://pirsa.org/17090051

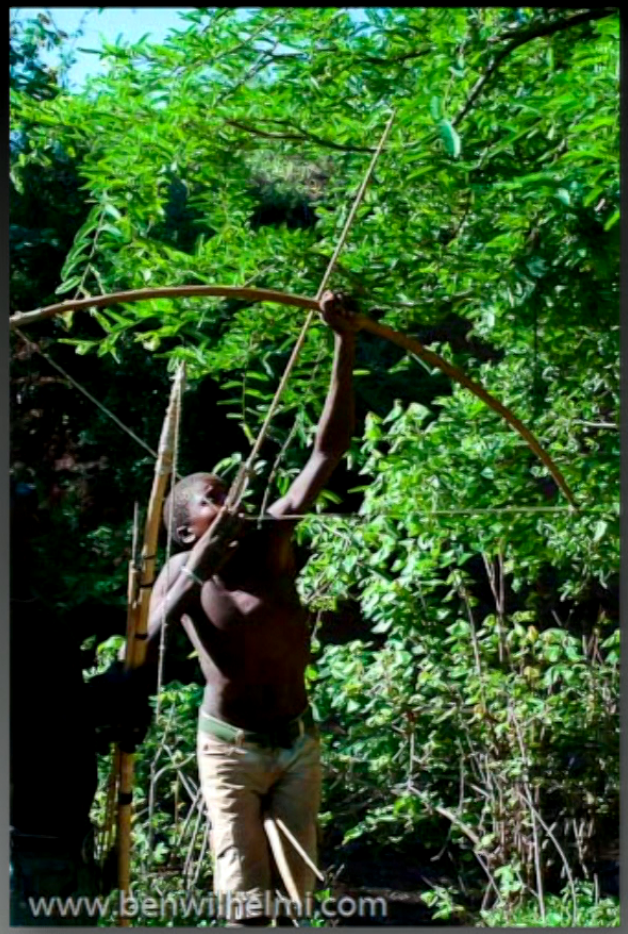
Abstract: I will present several studies showing a surprising pattern. Not only can preschoolers learn abstract higher-order principles from data, but younger learners are actually better at inferring unusual or unlikely principles than older learners and adults. This pattern also holds for children in Peru and in Headstart programs in Oakland, California. I relate this pattern to computational ideas about search and sampling, to evolutionary ideas about human life history, and to neuroscience findings about the negative effects of frontal control on wide exploration. My hypothesis is that our distinctively long, protected human childhood allows an early period of broad hypothesis search, exploration and creativity, before the demands of goal-directed action set in.

Life History and Learning:

AKA A Grand Unified Theory of Childhood

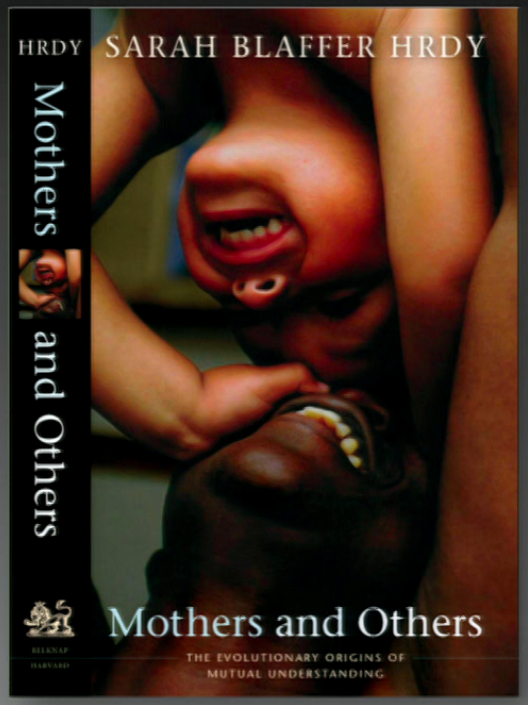
Alison Gopnik

Life History and The Evolution of Childhood


www.benwilhelmi.com

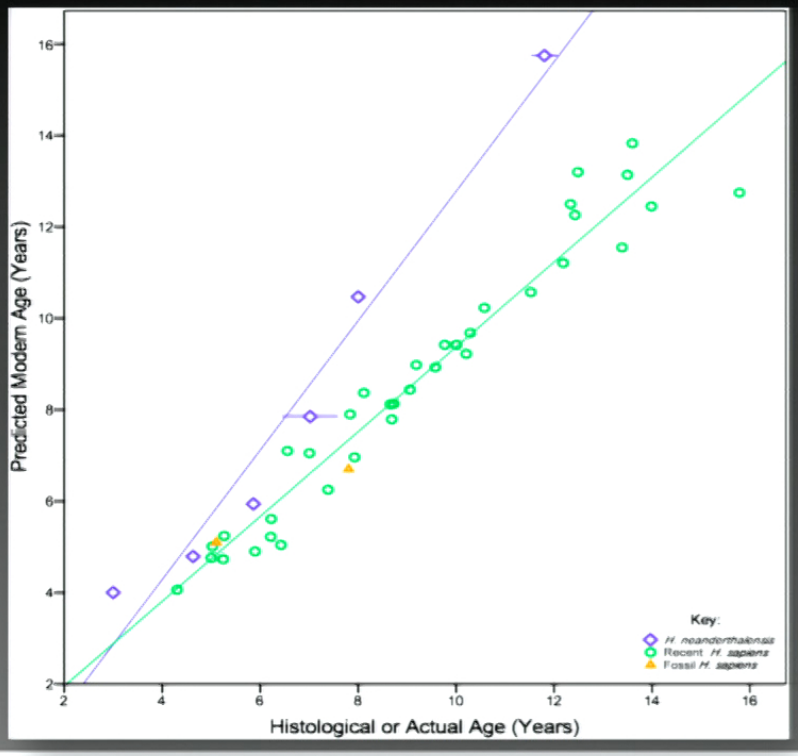
The Investment Triple Threat: Pair Bonding and Paternal Investment

The Triple Threat: Alloparents



The Triple Threat: Grandmothers

Fossil Evidence: Life History in Hominin Evolution

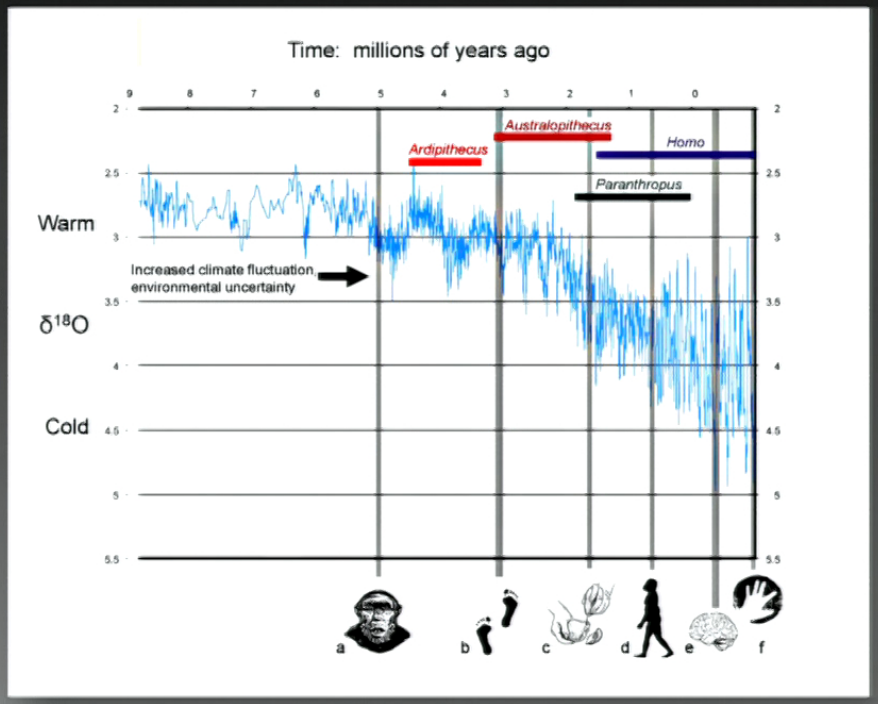


Smith T M et al. PNAS 2010;107:20923-20928

©2010 by National Academy of Sciences

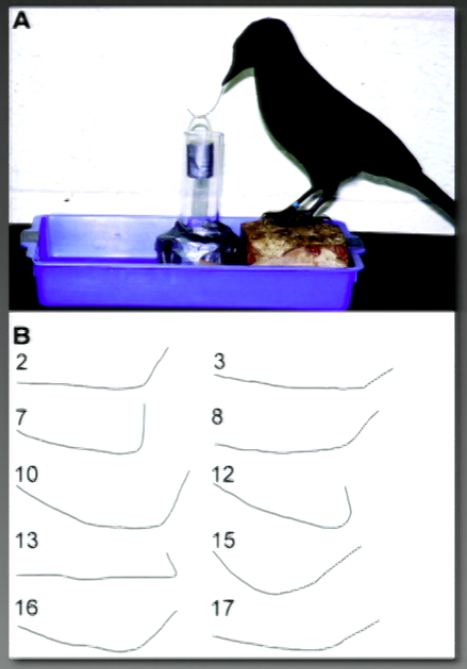
PNAS

Ecological Evidence: Climate Change Induced Humans

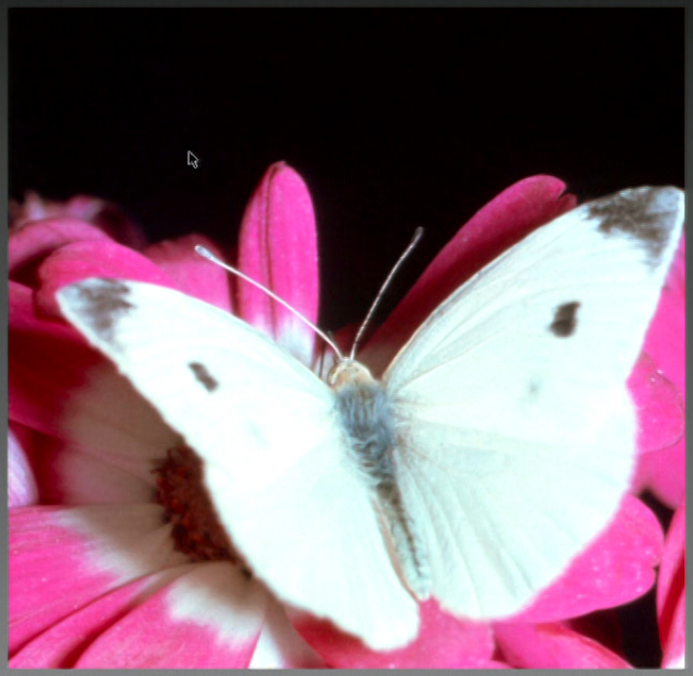


Comparative Evidence: Life History, Brain Size and Learning: Marsupials: Quokka vs Possums

Crows vs Chickens



Insects: Cabbage Whites & Kale



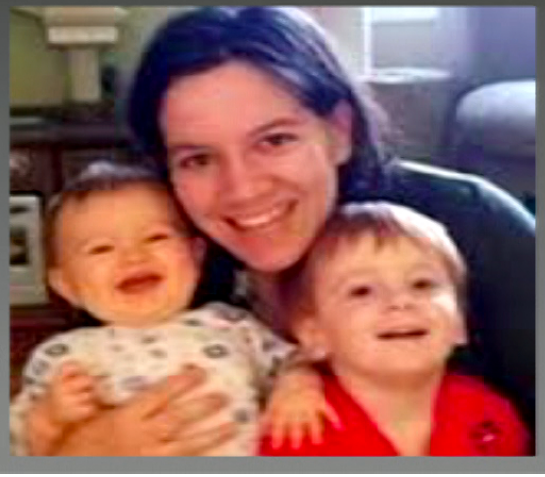
THE THEORY THEORY 2.0

Learning Generative Probabilistic Bayesian Models from Statistical Data

- Gopnik & Wellman, Psychological Bulletin, 2012
- Gopnik, Science, 2012

The Sampling Hypothesis

E. Bonawitz, S. Denison, T. Griffiths, A. Gopnik. (2014). Probabilistic Models, Learning Algorithms, Response Variability: Sampling in Cognitive Development. *Trends in Cognitive Sciences* doi.org/10.1016/j.tics.2014.06.006



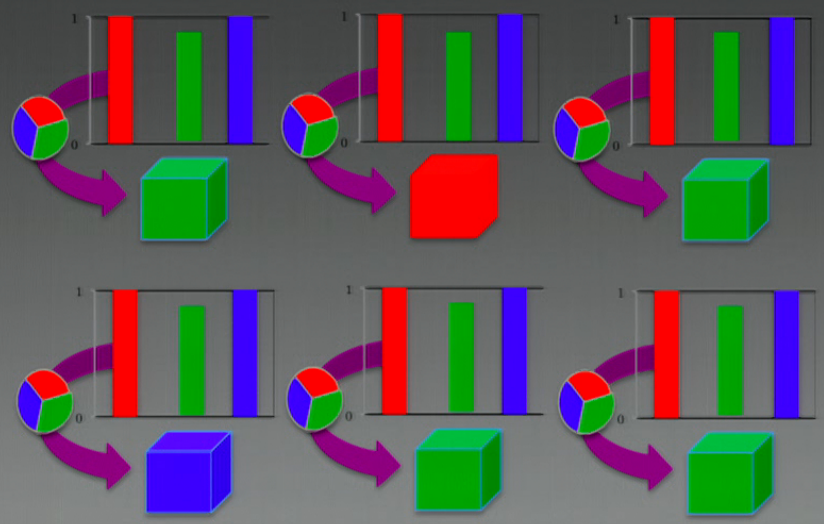
The Sampling Hypothesis

Posterior Distribution

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h' \in \mathcal{H}} P(d|h')P(h')}$$

Problem:

Large Hypothesis Space



Solution:

Monte Carlo Methods

- **Variable–hypotheses sampled "at random"**
- **Systematic—more probable hypotheses sampled more often.**

Exploitation vs. Exploration

Low-temperature search

Quick to settle on high-probability answer

May miss low-probability answer

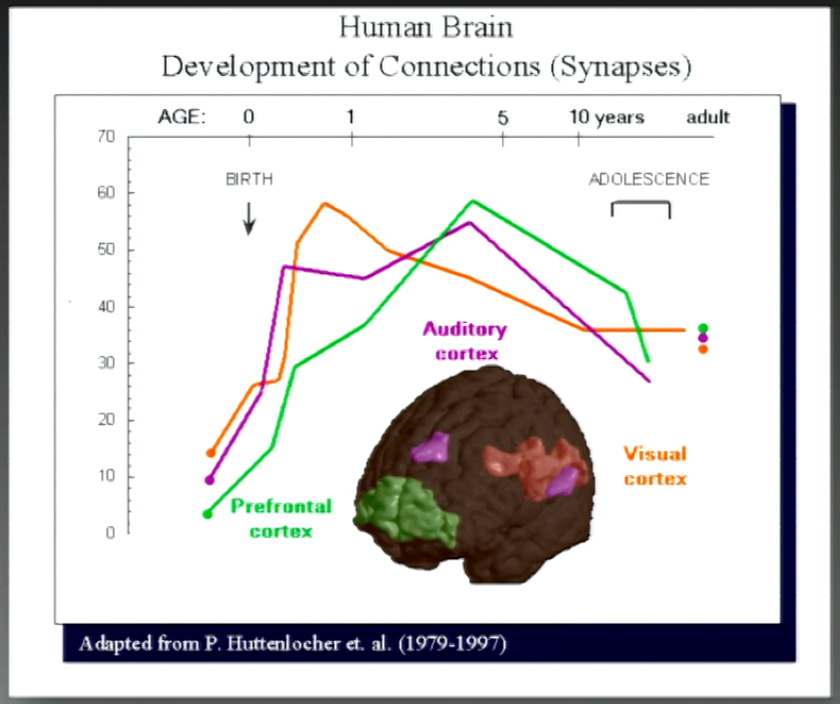
High-temperature search

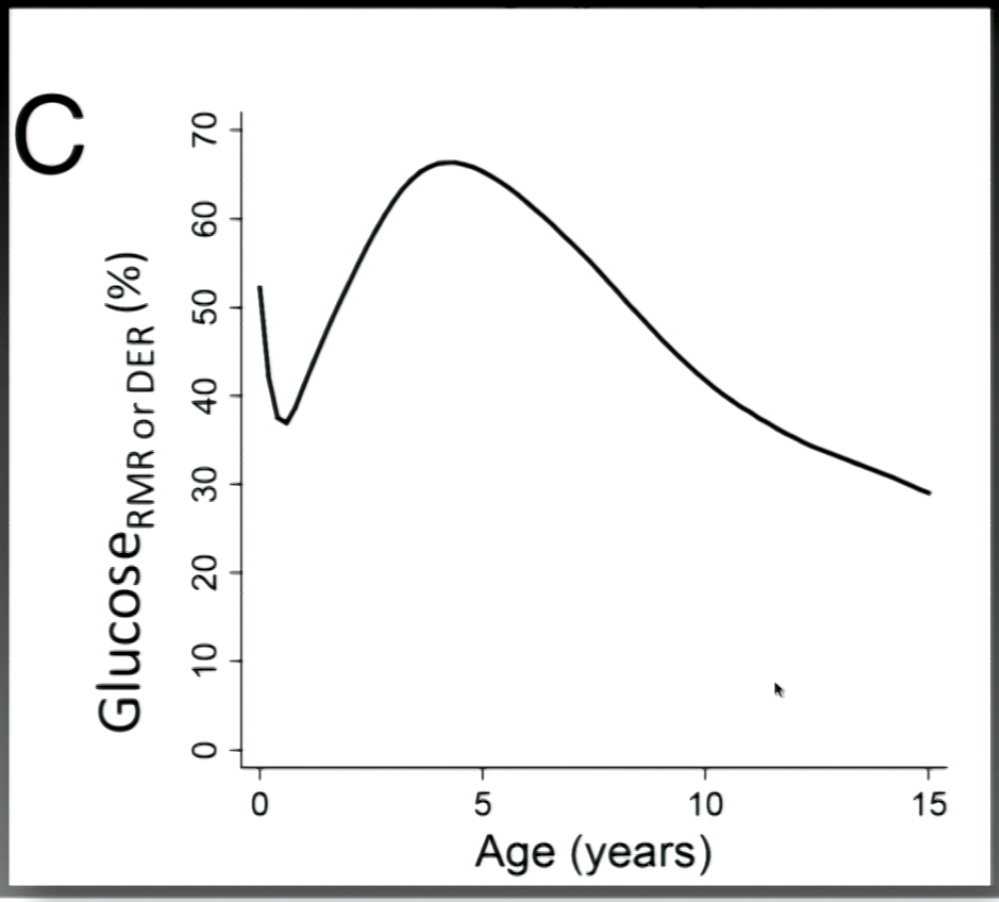
Slow to settle on high-probability answer

More likely to find low-probability answer

Simulated Annealing

Evidence From Neuroscience





Christopher W. Kuzawa et al. PNAS 2014;111:13010-13015

©2014 by National Academy of Sciences

PNAS

Less Frontal Control Leads to More Exploration

- Thompson-Schill et al. Current Directions in Psychological Science. 2009
- Jazz Improvisers Show Less Frontal Activity. Limb et al.
- Frontal TDCS Leads to More Divergent Thinking. Chryskiou et al. Cognitive Neuroscience 2013
- Transient Frontal Release Is Associated with Learning. Bassett et al. PNAS 2011

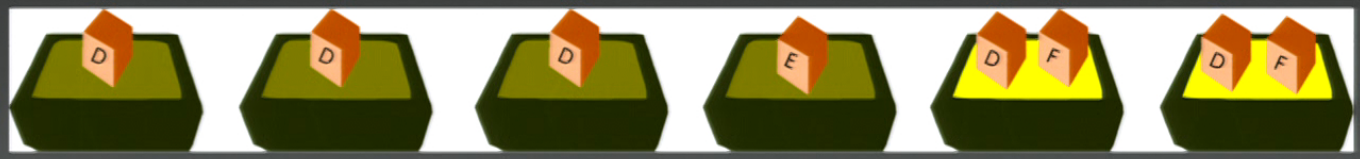
The Blicket Detector

Some blocks are blickets. Blickets make the blicket detector light up and play music.

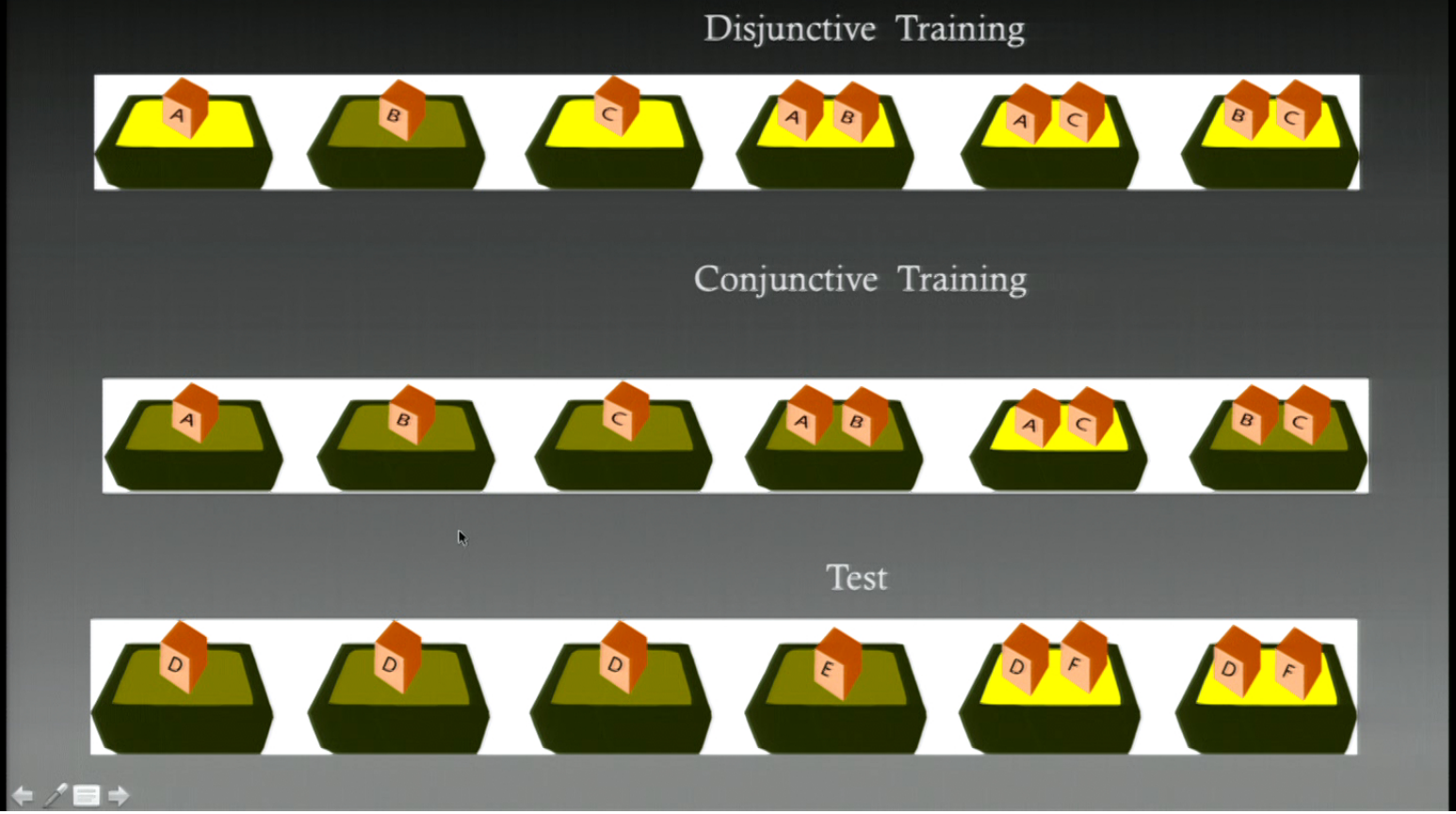
When Younger Learners are More Open-Minded

- A. Gopnik, T. Griffiths, & C. Lucas (2015). *Current Directions in Psychological Science*, *24* (2), 87-92
- C. Lucas, S. Bridgers, T. Griffiths, & A. Gopnik (2014). *Cognition*. 131, 2, 284–299.
- A. Gopnik, S. O'Grady, C. Lucas, T. Griffiths A. Wente, S. Bridgers, R. Aboody, H. Fung, R. E. Dahl, (2017). *PNAS*.

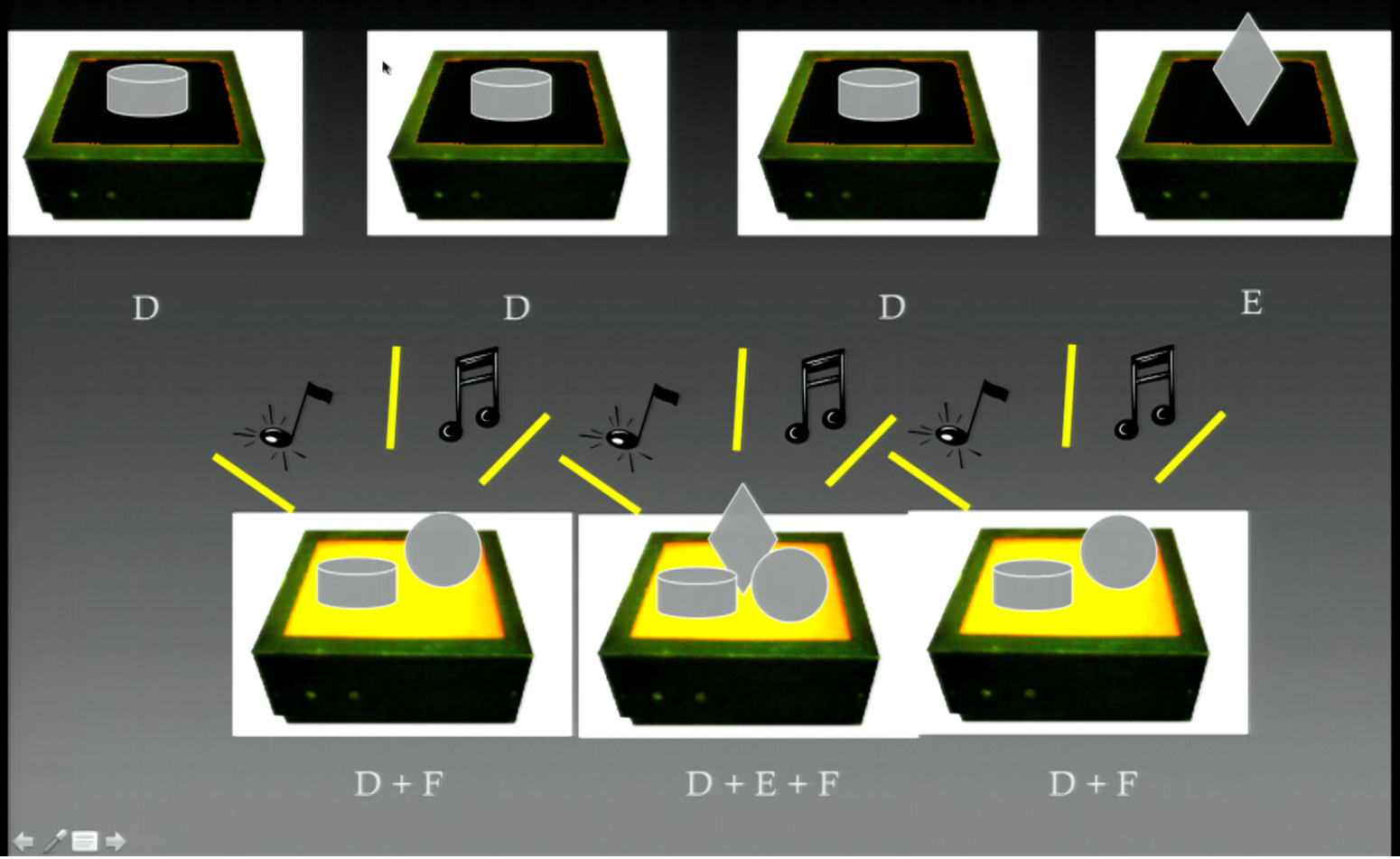
Which objects are blickets?

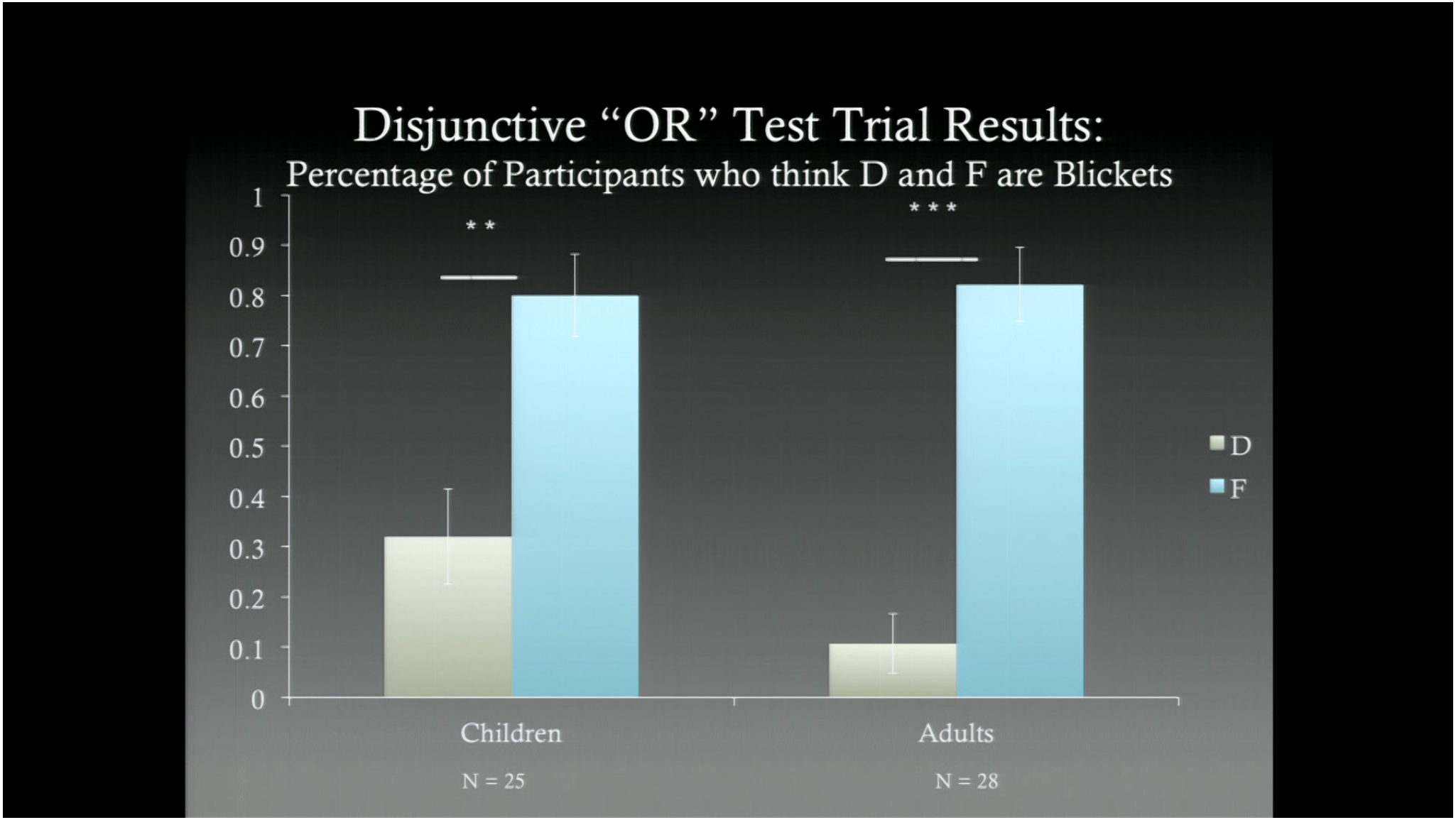


Is D a blicket? Is E a blicket? Is F a blicket?

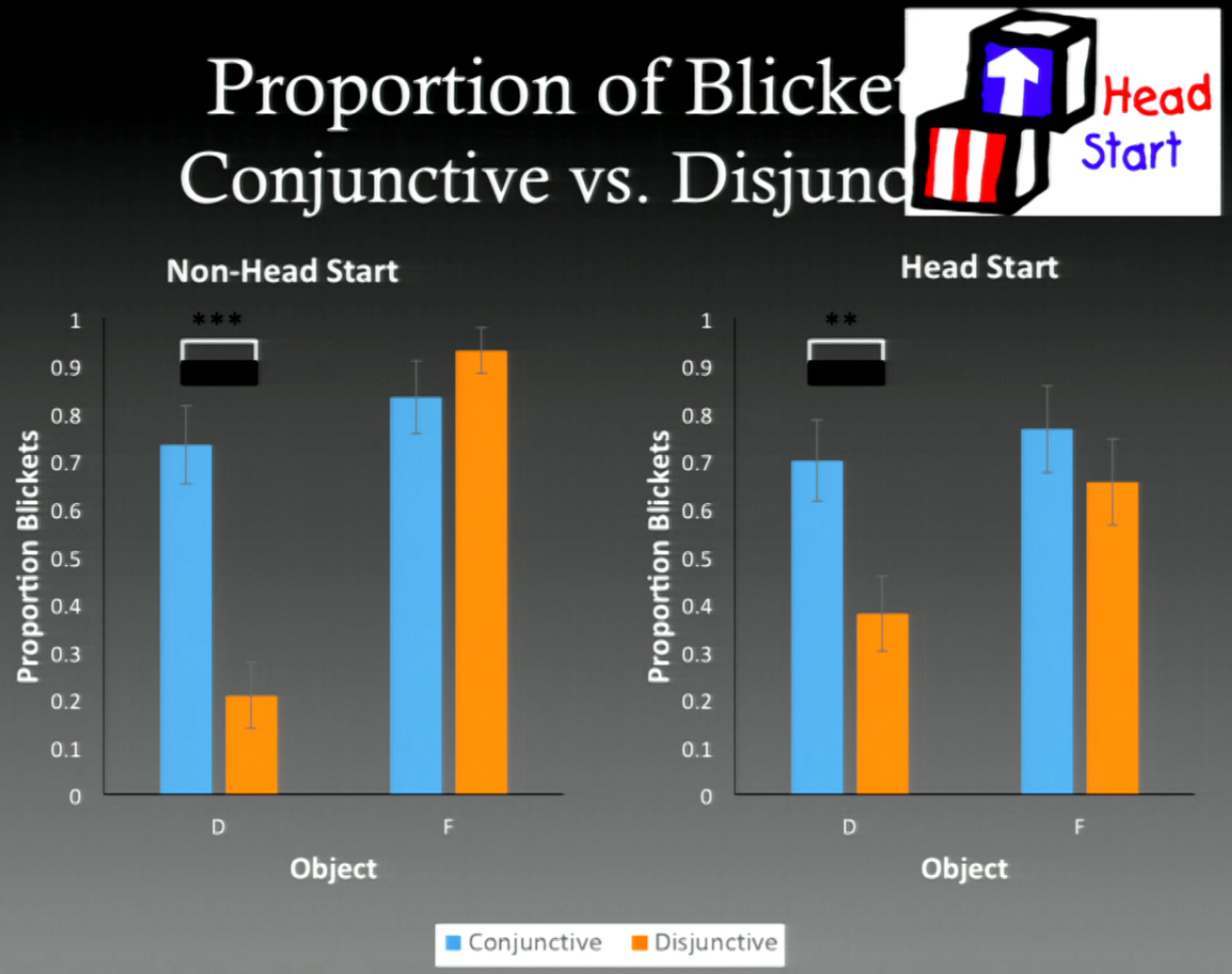

Disjunctive Training
A
B
C
A B
A C
B C
Conjunctive Training
A
B
C
A B
A C
B C
Test
D
D
D
E
D F
D F

Functional Form Procedure: Test trial in Experiment 2




Disjunctive "OR" Test Trial Results:
Percentage of Participants who think D and F are Blickets
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
* *
* * *
D
F
Children
N = 25
Adults
N = 28

Proportion of Blicke
Conjunctive vs. Disjunc



Is Causal Learning WEIRD? Innova Schools in Peru

- Serve Peru's emerging middle-class, many of whom are first- or second-generation immigrants from the Andes or the Amazon. N - 60
- The average gross income for families is roughly $1,200, of which 20-25% of their budget is spent on education
- Compared to Peru undergraduates, N= 83.

Change Across Development: Gopnik et al. PNAS, 2017.

- Gradual change with accumulated experience?
- Discontinuities in life history transitions? Preschool to school age, school age to adolescence, adolescence to adulthood.

Causal Inference and Social Theories: Seiver, Gopnik & Goodman, Child Development 2013

- Trait vs. situation attribution
- Trait concepts only appear explicitly and simultaneously around age 6
- Kelley and co-variation

	Skateboard	Scooter
Josie	Play	
Sally		

DOLL CONDITION	Skateboard	Scooter
Josie	Play, Play, No P, Play	Play, Play, No P, Play
Sally	No P, No P, Play, No P	No P, No P, Play, No P

DOLL CONDITION	Skateboard	Scooter
Josie	Play, Play, No P, Play	Play, Play, No P, Play
Sally	No P, No P, Play, No P	No P, No P, Play, No P

CONTROL CONDITION	Skateboard	Scooter
Josie	Play, Play, No P, Play	
Sally		No P, No P, Play, No P

Internal Explanations

Age

Size

Mental States

Emotions

Experience

The scooter is fun!

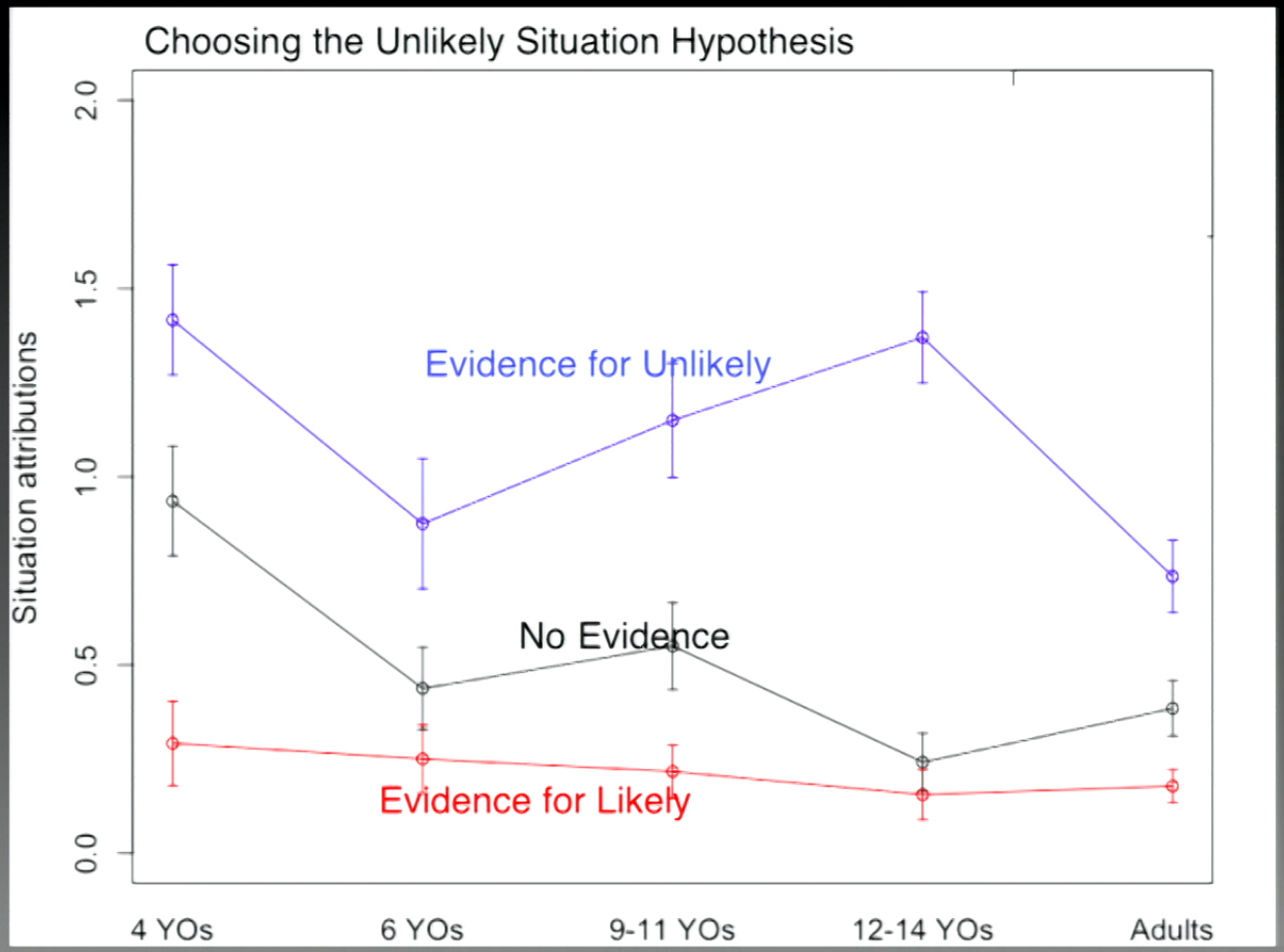
The skateboard is wobbly.

External Explanations

Activity

Situation

Friend


Choosing the Unlikely Situation Hypothesis
Situation attributions
0.0
0.5
1.0
1.5
2.0
Evidence for Unlikely
No Evidence
Evidence for Likely
4 YOs
6 YOs
9-11 YOs
12-14 YOs
Adults

Other Examples Where Younger Learners Are More Exploratory

Learning situation vs. trait attributions: Seiver et al. Child Development 2013

Learning relational vs. individual concepts: Walker et al. Cognition 2016

Searching a Semantic Network. Bonawitz et al. SRCD, 2015.

Learning foreign language vs. own language distinctions (Kuhl 2004, Werker, 2015).

Learning novel vs. familiar artifact uses. (German & Defeyter, 2000).