Understanding infants' curiosity-based learning: empirical and computational approaches

Katherine Twomey, Gert Westermann,
Chen Yu, Pierre-Yves Oudeyer

ICIS 2016

Understanding infants' curiosity-based learning: empirical and computational approaches

1) Infants' information seeking in a category exploration task.
Katherine E. Twomey, Ben Malem & Gert Westermann

2) A computational model of infants' curiosity-based learning.
Gert Westermann & Katherine E. Twomey

3) Social basis of sustained attention and exploration: Coordinated attention with parents trains infants' sustained attention skills.
Chen Yu & Linda Smith

4) The learning progress hypothesis: Theory and models of curiosity-driven exploration and its impact on development.
Pierre-Yves Oudeyer

Infants' information seeking in a category exploration task

Katherine E. Twomey
k.twomey@lancaster.ac.uk

Ben Malem
b.malem@lancaster.ac.uk

Gert Westermann
g.westermann@lancaster.ac.uk

LuCiD
The ESRC International Centre for Language and Communicative Development

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Babylab

Lancaster University

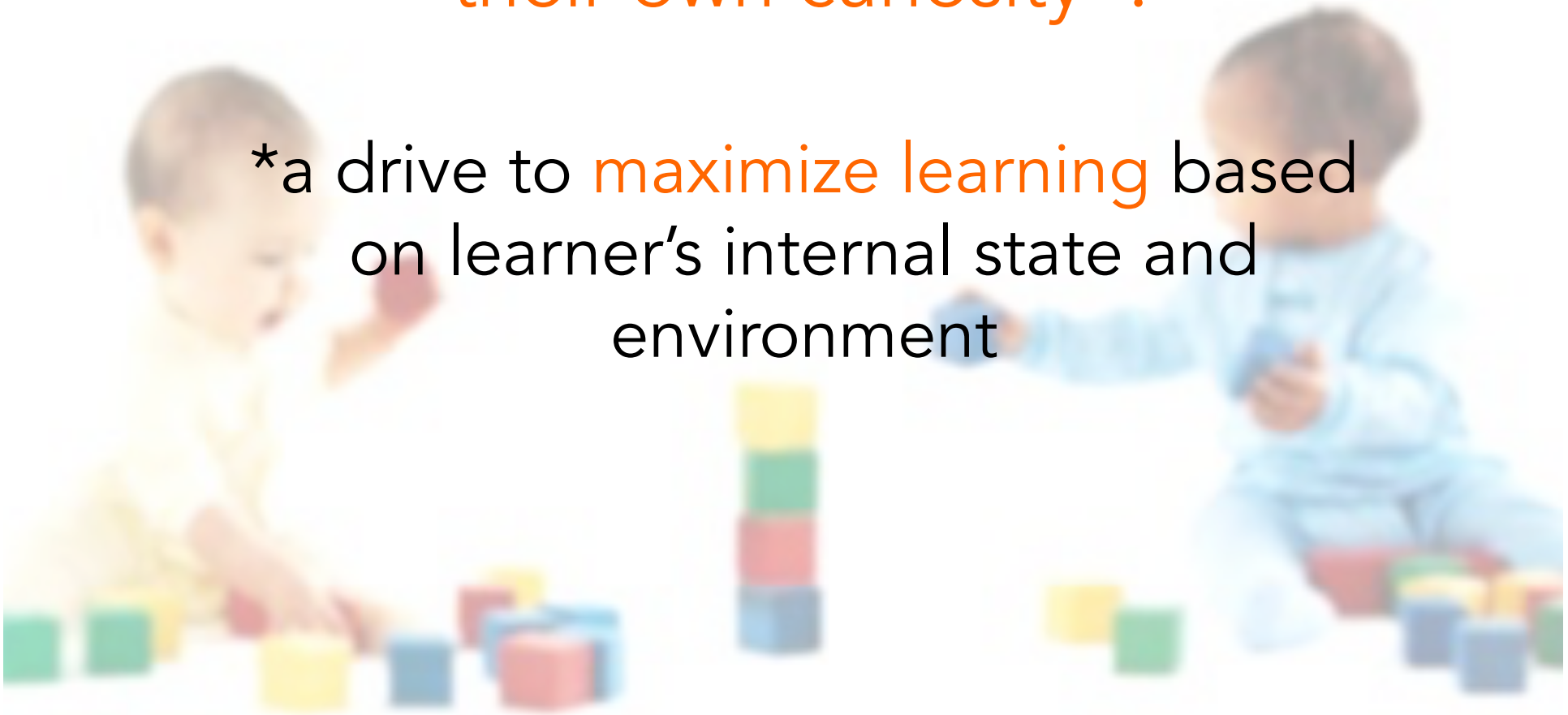

Order, duration
determined
a priori

Infants are curious learners!

Everyday learning proceeds via exploration

How do infants sample their learning environment based on their own curiosity*?

*a drive to maximize learning based on learner's internal state and environment



For a discussion see Kidd & Hayden (2015, *Neuron)*

1) Do infants select information systematically?

Kidd, Piantadosi & Aslin
(2012, *Plos One*; 2014, *Child Dev*)

- 7-8mos, looking time task
- Infants looked for longer at events with intermediate predictability

Twomey & Westermann (2015,. *Proc. ICDL-EPIROB*)

- Connectionist model which chose its own stimuli
- Suggested infants will switch between low and high complexity stimuli

2) If so, what level of complexity will infants generate? Does this interact with labeling?

Maximum? Mather & Plunkett (2011; *Cognition*)

- 10mos categorization
- maximum Euclidean distance

Intermediate? Twomey, Ranson & Horst (2014; *Infant Child Dev*)

- 30mos, categorization/word learning
  - medium perceptual variability

Minimum? Bulf, Johnson & Valenza (2011; *Cognition*)

- Newborns, visual sequence learning
- minimum unpredictability

Need a task that lets us retain experimental control but allow exploration

Quantify difficulty - perceptual distance

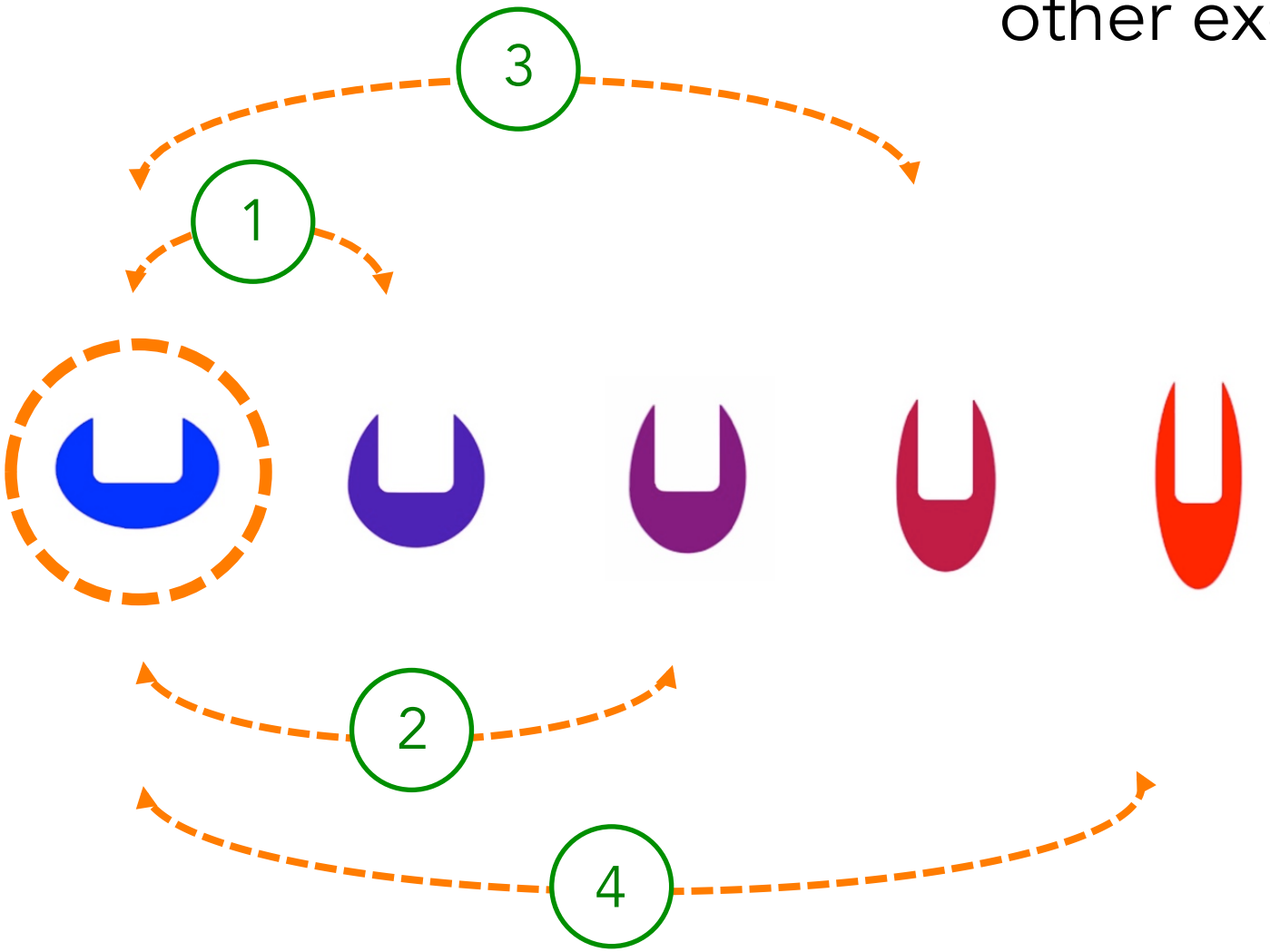
(cf. Mather & Plunkett, 2011; *Cognition*)

Visual stimuli that differ systematically

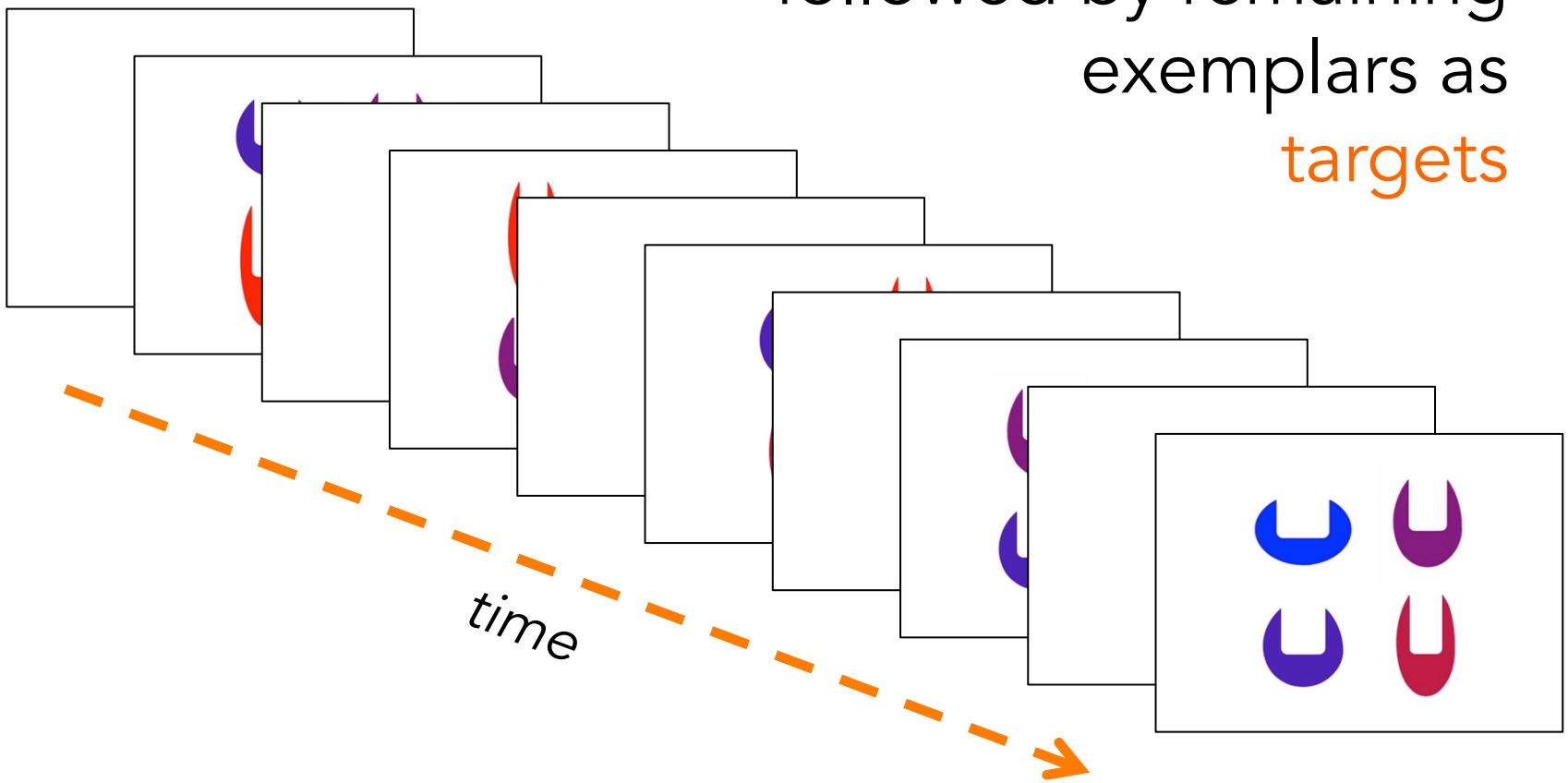
(cf. Althaus & Westermann, 2016; *JECP*)

A B C D E

Complexity: target distance along continuum from other exemplars



Control complexity: use each exemplar as a prime item once, followed by remaining exemplars as targets



Participants:

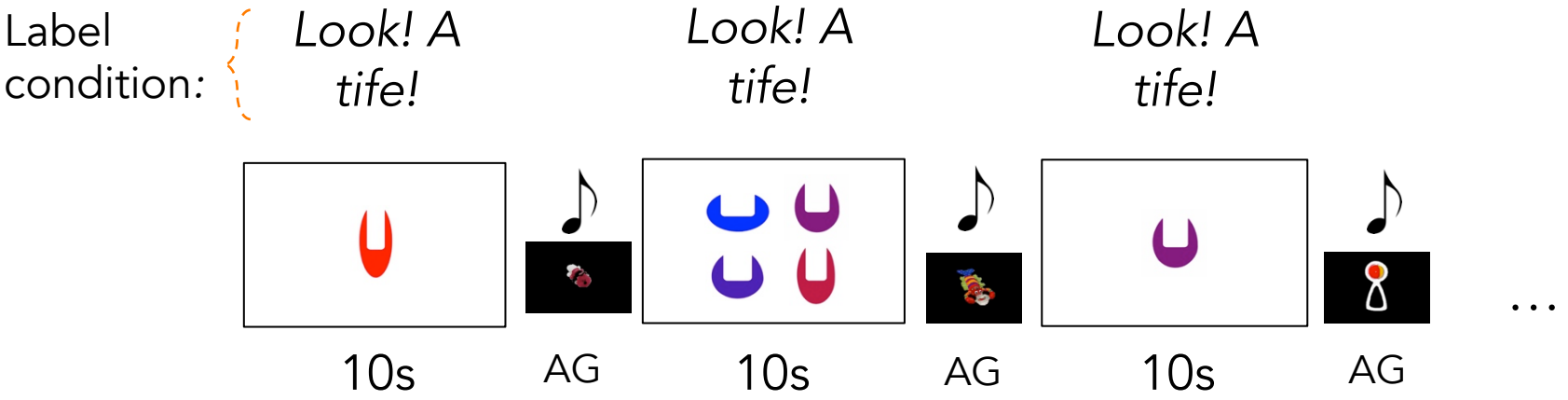
40 12mo infants

Procedure:

2 conditions: label (n = 20) / no-label (n = 20)

Static images on screen, eyetracked

Design:

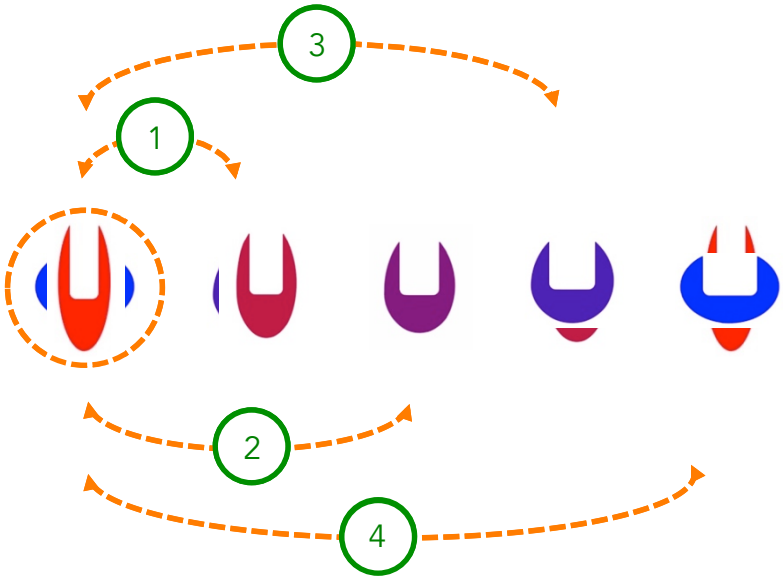


5 x prime-target pairs

Analysis

DV: looking to targets after peripheral primes

Equal chance of selecting each distance

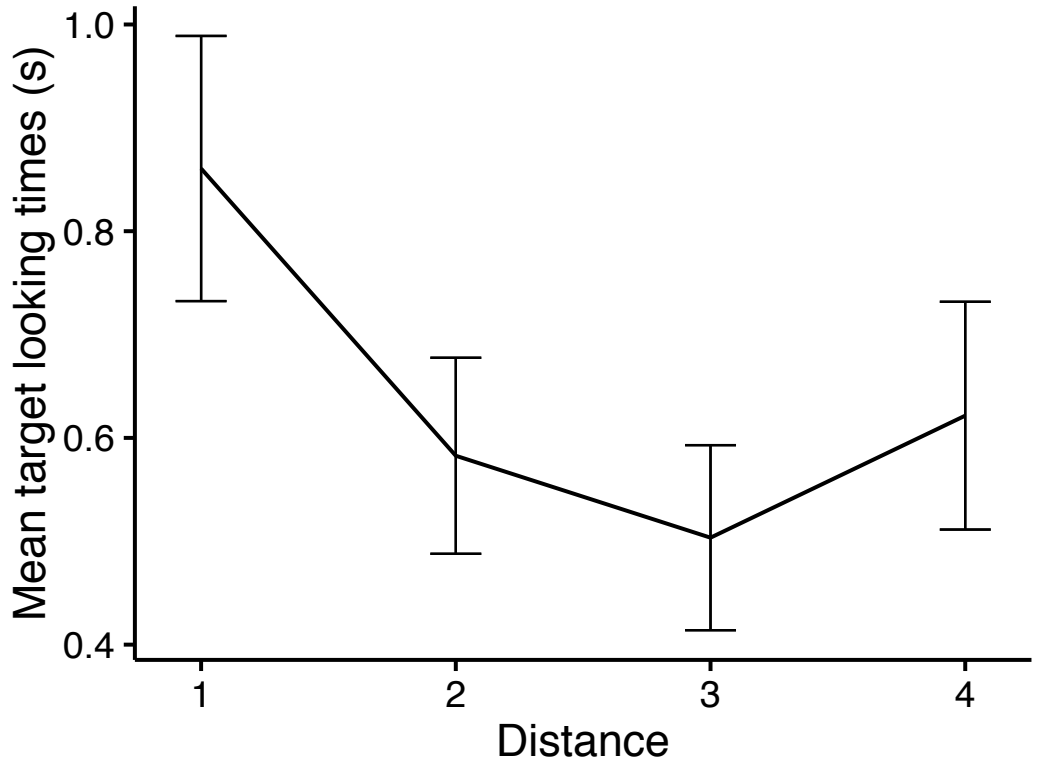


Fixed effects:
distance, label, interaction

Random effects: random intercepts for participant and target (Barr, Levy & Scheepers, 2013; *JML*)

Overall, targets that are closer to the primes elicit higher looking times

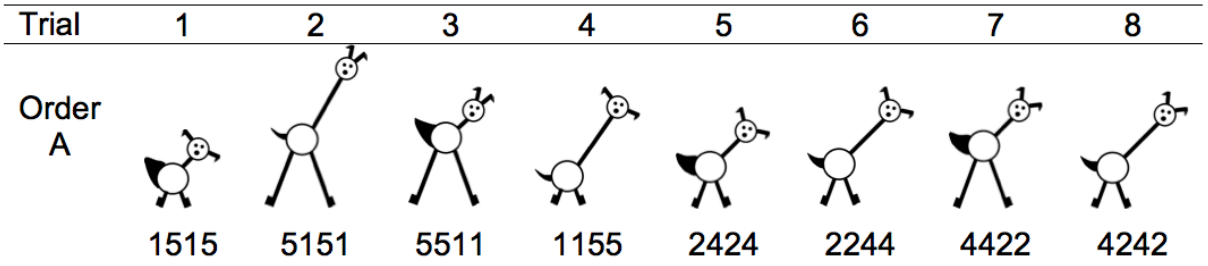
Main effect of distance: β = -0.4517, $\chi^2(1) = 21.02$, $p < .001$



Error bars represent 95% CIs

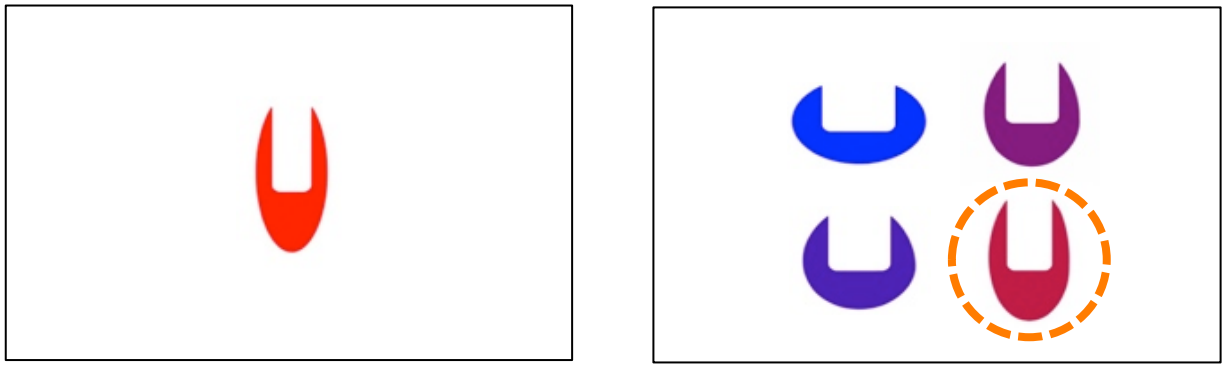
However: Twomey & Westermann (2015) analysed sequences of stimuli

Apparent "intermediate complexity" emerged from switching between stimuli of maximum and minimum complexity

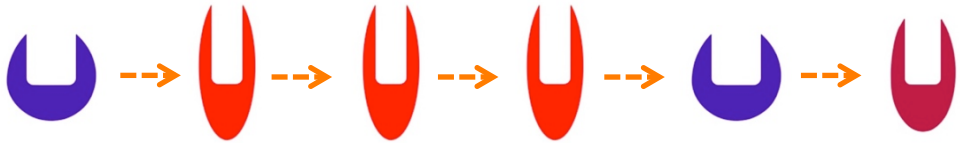


Is this really complexity minimization?

If so, infants should look at the least distant stimulus only

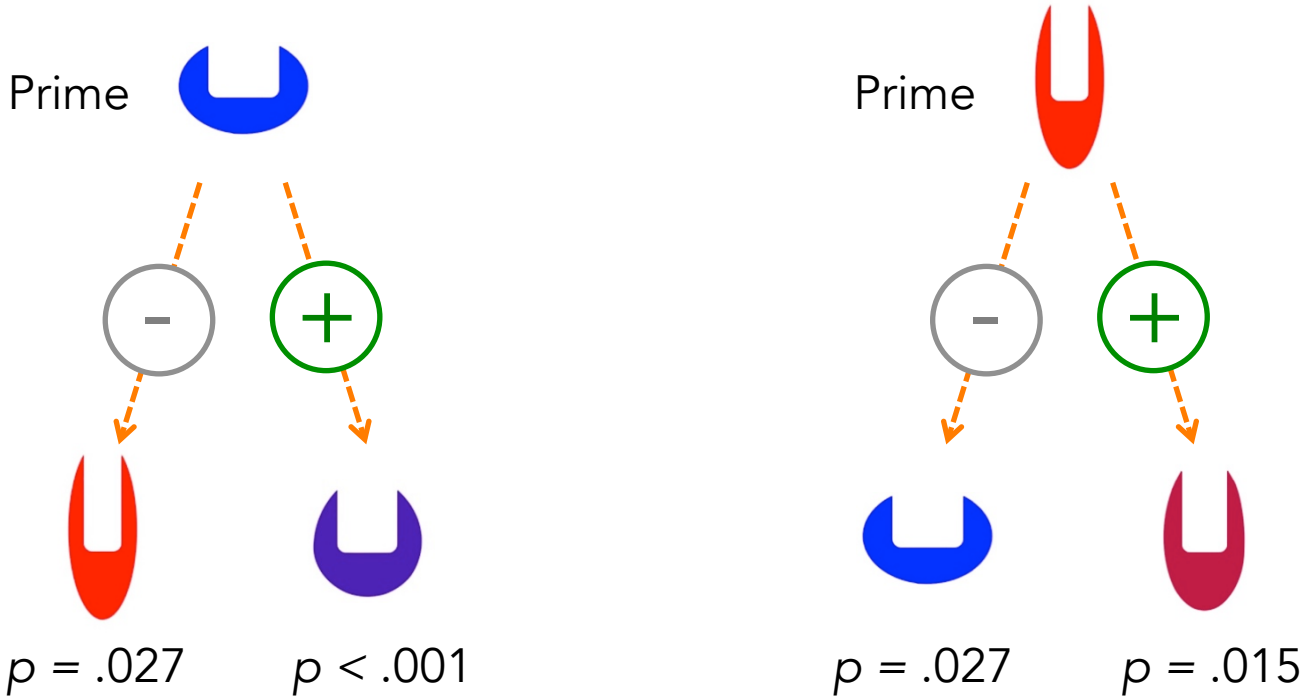


Track individual fixations



to generate exploratory sequences

Where do infants look first?



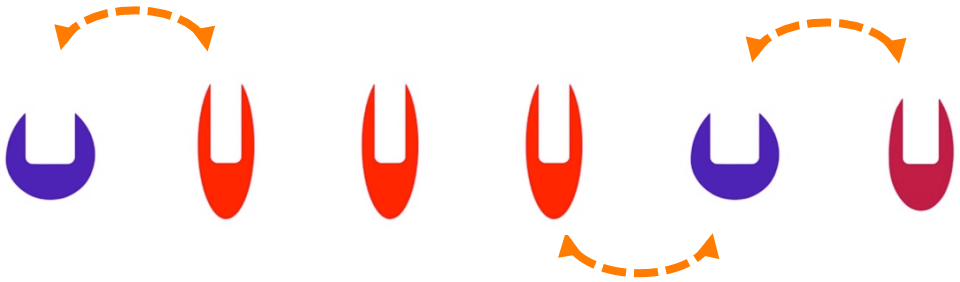
$p = .027$ $p < .001$ $p = .027$ $p = .015$

Initially, infants look at the exemplar with the shortest distance from the prime:

minimize complexity

exact binomial tests, chance = 0.25

For each sequence,
record transition between targets

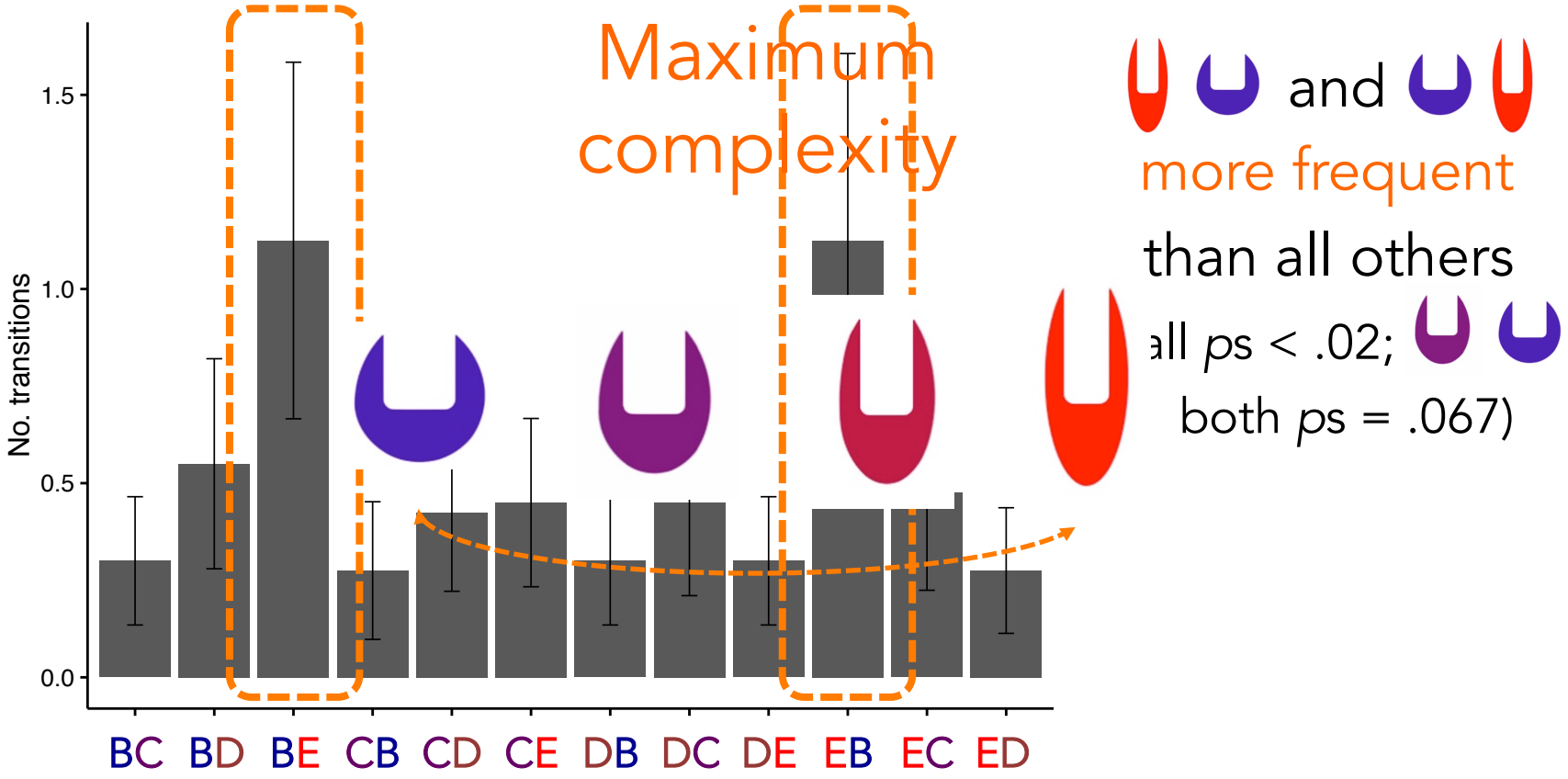


CE x 1 (distance of 2)

EC x 1 (distance of 2)

CD x 1 (distance of 1)

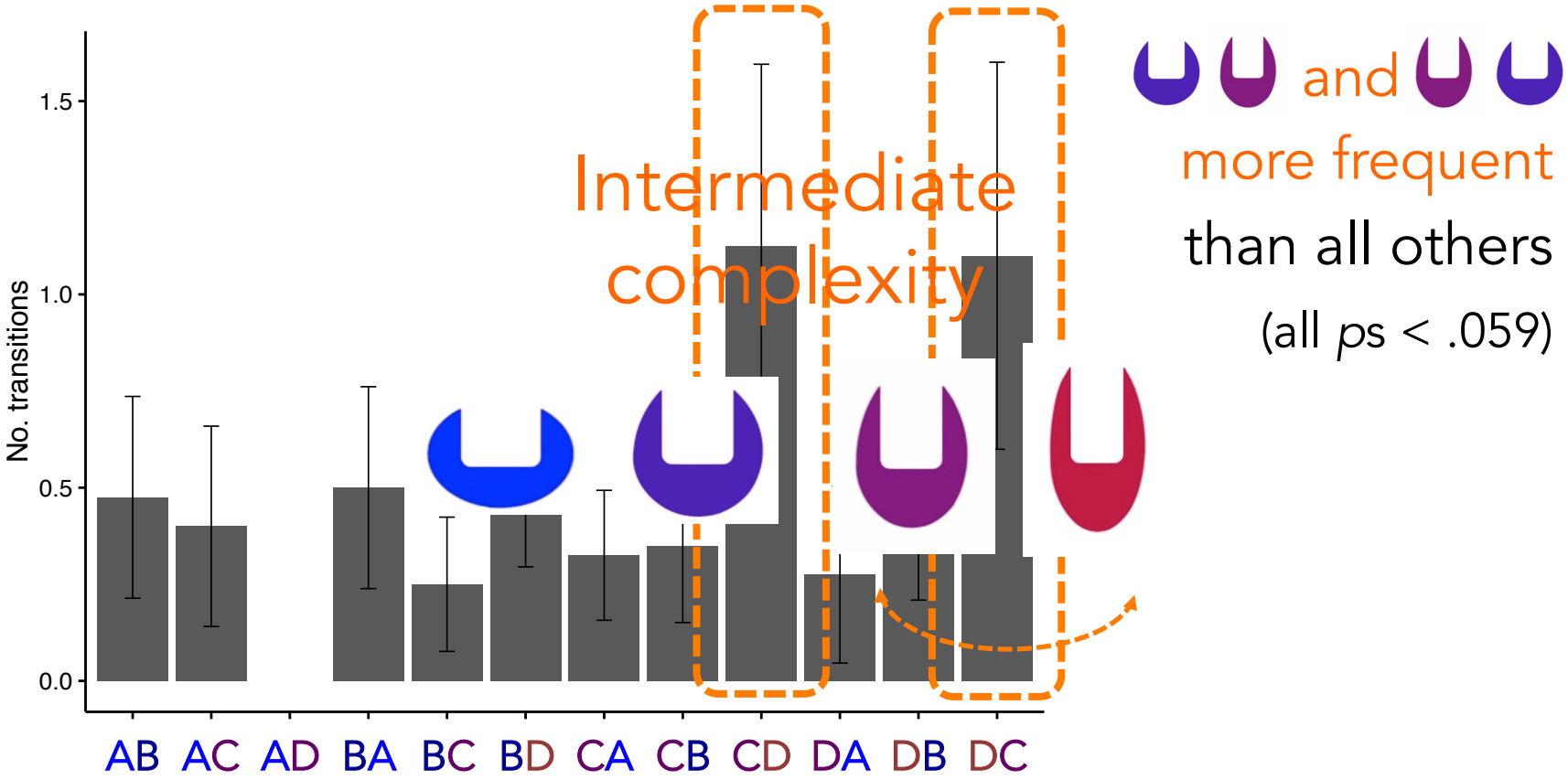
Which transitions are most common after ∪ ?



Main effect of transition, $\chi^2(11) = 56.87$, $p < .001$

Error bars represent 95% CIs

Which transitions are most common after  ?



Main effect of transition, $\chi^2(11) = 59.10$, $p < .001$

Error bars represent 95% CIs

Although looking times suggest infants minimize complexity overall, transitions indicate that in real time, infants generate intermediate or maximal complexity

What patterns of looking do infants show?

Switching

(Kovack-Lesh, Horst & Oakes, 2008; *Infancy*)

51/80 sequences showed switching

(minimum criterion: look away then return, e.g., ABA)

BUT: complexity of switch depends on prime

80% of switches: distance of 1

63% of switches: distance of 3

What's the story?

1) Do infants select information systematically?

Yes!

- overall looking – longer looking to smaller distances
- first look – to smallest distance
- transitions – systematically maximized or minimized transition distances

2) If so, what level of complexity will infants generate?

While overall looking times suggested a preference for less complex stimuli, fine-grained analyses revealed patterns of switching that generated intermediate or maximum complexity (for a discussion of temporally-based analyses of infant looking see Balas & Oakes, 2015; *Proc. ICDL-EPIROB*)

Implications

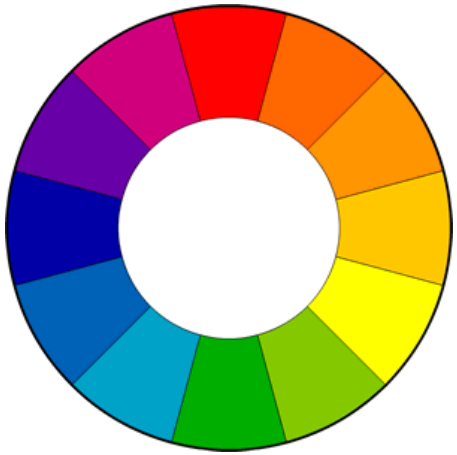
Comparison is important: simultaneous stimulus presentation leads to better category learning than successive presentation (Oakes et al., 2009; *JECP*)

Transitions are important: infants who see stimuli presented in orders which maximize transitional complexity learn best (Mather & Plunkett, 2011; *Cognition*)

Highlights the importance of switching as a mechanism of information selection (see Kovack-Lesh, Oakes & McMurray, 2012; *Infancy*)

But transition preference was context dependent (Kovack-Lesh et al., 2012): curiosity-driven information selection depends on interaction between learner's internal state and environment

Challenges



Why did switching differ by prime?

- Design stimuli from a category without obvious boundaries

No effect of label

- Test in older children, adults

New paradigms

- Selection without replacement – gaze contingency

Theory development

- Is complexity objective? Subjective? Novelty? Predictability?

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Leverhulme Trust

BRITISH ACADEMY for the humanities and social sciences

Babylab | Lancaster University

LuCiD
The ESRC International Centre for Language and Communicative Development

Questions?

"We lack even the most basic integrative theory of the basis, mechanisms, and purpose of curiosity"

For a discussion see Kidd & Hayden, 2015

Questions

Mechanism?

- Information gap, triggers info seeking ? (Loewnestein, 1994; Twomey & Westermann, under review)?
- Novelty maximisation [but: familiarity preference]?
- Uncertainty minimization? (Oudeyer & Kaplan, 2007).
- Understanding causality?

How does curiosity interact with environment/context? (Baranes et al. 2014)

How do we define novelty? (Mather 2013) How do we define complexity? Objective? Subjective?

How can we differentiate empirically between these mechanisms?