UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Infants Use Social Signals to Learn from Unfamiliar Referential Cues

Permalink
https://escholarship.org/uc/item/0n76z4dn

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Wu, Rachel
Kirkham, Natasha
Swan, Kristen
et al.

Publication Date
2011

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Infants Use Social Signals to Learn from Unfamiliar Referential Cues

Rachel Wu (r.wu@bbk.ac.uk), Natasha Z. Kirkham (n.kirkham@bbk.ac.uk), Kristen A. Swan (k.swan@bbk.ac.uk), Teodora Gliga (t.gliga@bbk.ac.uk)

Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London
Malet Street, London, WC1E 7HX, UK

Abstract

Infants are bombarded with a bewildering array of events to learn. In such an environment, referential cues (e.g., gestures or symbols) highlight which events infants should learn. Although many studies have documented *which* referential cues guide attention and learning during infancy, few have investigated *how* this learning occurs. The present eye-tracking study provides clear evidence for a social scaffolding process: When preceded repeatedly by communicative signals (i.e., a face addressing the infant), 9-month-olds learned that arbitrary cues predicted the appearance of an audio-visual event. Importantly, the arbitrary cues continued to guide learning of these events, even after the face disappeared from the screen. A control condition confirmed that learning from arbitrary cues alone was unsuccessful, and that eventual success was not just due to extended practice. These results are discussed in terms of a theory of cue scaffolding.

Keywords: attention cues; multimodal binding; infant eye-tracking; cognitive development; social cues.

Introduction

Human adults use referential cues, such as gaze, pointing, and arrows, to direct attention and communicate information (Kita, 2002). By the first year, infants orient to the direction of gaze and pointing fingers (e.g., Gredebäck, Melinder, & Daum, in press; Hood, Willen, & Driver, 1998; Senju & Csibra, 2008) and, more importantly, expect these cues to communicate about something in the indicated location (Gliga & Csibra, 2009; Tomasello, Carpenter, & Liszkowski, 2007). It is not until three years of age, however, that children can robustly use more arbitrary cues, such as arrows, to find hidden objects (Leekam, Solomon, & Teoh, 2010). How do we learn the communicative intent underlying novel referential cues in order to use these cues to learn about the world?

A number of studies have highlighted the importance of ostensive-communicative signals (e.g., infant-directed speech, eye contact, smiling) for the efficacy of referential cues and for better learning during infancy. These signals are necessary for referential gaze shifts to successfully orient attention in 4- and 6-month-old infants (Senju & Csibra, 2008; Farroni, Mansfield, Lai, & Johnson, 2003). These signals also facilitate learning from referential cues such as gaze (Wu, Gopnik, Richardson, & Kirkham, 2010; Wu & Kirkham, 2010) and pointing (Yoon, Johnson, & Csibra, 2008). Though these studies describe the importance of pairing communicative cues with familiar cues, they do not show whether communicative cues can scaffold learning from novel cues during infancy. Leekam et al. (2010) showed that 3-year-olds could use a replica cue (e.g., a miniature version of a target container) to find stickers hidden underneath the target container, but only if the replica cue was presented with an engaging face. It is therefore possible that the initial pairing of ostensive-communicative cues and novel cues may be essential for infants to learn the communicative message behind an unfamiliar cue, leading to enhanced learning about the cued events.

One way to test this hypothesis with infants is to use a paradigm involving cues that can, on their own, successfully orient attention but do not produce optimal learning about the target of attention. For example, a bright light or loud noise might drag attention to a particular location, without communicating anything beyond itself. Indeed, in a recent study when arbitrary attention cues were used, although 8-month-olds' attention was oriented successfully, learning of the cued event was poor (Wu & Kirkham, 2010). The present study investigated whether social communicative signals can scaffold 8-month-olds' learning of multimodal events from a novel attention cue (e.g., bright red flashing squares that surround the location of a target audio-visual event). In order to test this scaffolding hypothesis, the appearance of the flashing square cue was preceded by an engaging face that addressed the infant and verbally encouraged her to look at a specific event, but did not provide any directional information (*Social Scaffolding* condition). A recent study showed that infants are better at extracting statistical rules from sequences of non-social stimuli (e.g., tones) if they first heard those rules instantiated in social stimuli (i.e., speech; Marcus, Fernandes, & Johnson, 2007). Therefore, we made a further prediction that infants in the Social Scaffolding condition would continue learning from the novel cue even after the face was no longer present. To show this, at the end of the first part of the experiment (the Training phase), we presented infants with a second series of audio-visual events that were cued only with the flashing square (the Generalization phase). To ensure that practice with flashing squares could not account for the transfer learning effects, we presented another group of infants with only flashing square cues for both Training and Generalization phases (the *Extended Practice* condition). Importantly, this study is concerned with the *product* of attention – specific, accurate learning in a busy multimodal environment, rather than increased attention to the cued event. We predicted that when paired with social signals, novel flashing cues would elicit multimodal learning of cued rather than non-cued locations (similar to gaze cues paired with social signals, Wu & Kirkham, 2010). Furthermore, we predicted that infants can then continue learning from these novel cues even when they are not paired with social signals later. Without this initial pairing, we predicted that infants would only learn about the locations of cues (e.g., flashing square condition in Wu & Kirkham, 2010), and would show no transfer of learning.

Methods

Participants Thirty-one 8- to 9-month-old infants participated in one of two conditions: Social Scaffolding (N = 16, 5 girls, 11 boys, *M* = 8 months, 14 days, range: 7;24-9;12) or Extended Practice (N = 15, 9 girls, 6 boys, *M* = 8 months, 21 days, range: 7;29-9;21). Two infants were excluded from the final analyses (one from each condition) due to fussiness (i.e., completing only 1 out of 8 blocks). Infants were recruited via local-area advertisements and given t-shirts for participating.

Stimuli and Procedure Infants sat in a car seat 50 cm from a Tobii 1750 eye-tracker with a built-in 17” monitor, while their caregivers sat behind them. All infants were calibrated to at least 4 points before the experiment started. In both conditions, infants were shown up to 8 blocks of familiarization and test stimuli. The first four blocks consisted of the Training phase, while the second four blocks were the Generalization phase. During each phase infants were familiarized with two audio-visual events. These events were counterbalanced between participants with half of the infants seeing the audio-visual animations during Training that the other half saw during Generalization. Each block consisted of six familiarization trials and two test trials. For the Social Scaffolding condition, the familiarization trials in the first four blocks showed a central face that looked at the infant and said, “Hi baby, look at this!” This portion of the clip lasted 4 seconds. Then, the face froze with a smile while identical animations appeared in diagonally opposite corners within white frames. The animations during training were one of two sets: 1) two cats making *bloop* sounds and two buses making *whoosh* sounds or 2) two ducks making *brring* sounds and two dogs making *boing* sounds. A red flashing square was presented around the white frame in a lower corner, containing one of the animations. The trial ended 5 seconds later. Infants in the Extended Practice condition did not see the face during Training, only the flashing cues and audio-visual events. The Generalization phase of both conditions displayed the other two sets of audio-visual events with only the red flashing square cues (Figure 1). The test trials (5 s) were identical in both phases across both conditions and displayed four blank white frames (where objects previously had appeared), while playing sounds that were associated with a particular pair of objects. These blank test trials asked infants to predict where the associated objects would appear. Throughout an experimental session, two locations were cued, and within a trial, two locations contained objects that were paired with a particular sound (Figure 2). During familiarization, we were interested in how long infants looked to object and no-object locations (relative to the multimodal information) and whether this looking was mediated by novel flashing cues and social signals. During test trials, only auditory stimuli were played, so the locations were either matched or unmatched relative to the objects paired with the sound presented during familiarization. Stationary kaleidoscopic attention getters with ringing sounds looped after every trial until the infants re-engaged with the screen.

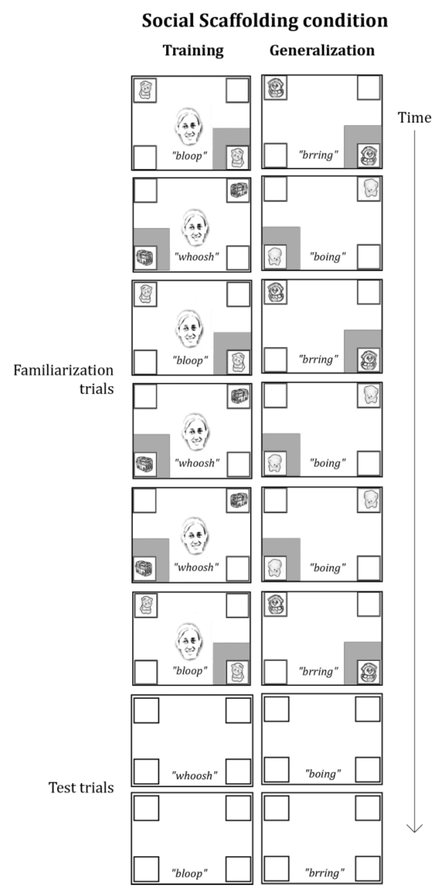

Figure 1: Schematic of one block of Familiarization and Test trials from the Social Scaffolding condition (Training and Generalization phases). The presentation of familiarization events was pseudo-randomized among infants, and test trial order was counterbalanced. All stimuli were in full color on a black background. The gray box around a frame represents a red flashing cue.

Data Reduction and Analyses Four regions of interest (ROIs) were manually delimited around the four corners where animations were played for both Familiarization and Test trials. We measured the accumulated looking time within each of these locations for the entire duration of every trial (see Figure 2). Only fixations longer than 100ms were considered in the final analyses. For each region of interest (ROI), we reported the proportional looking time, which was calculated in each trial for every infant by dividing the total looking time in that ROI by the total looking time in all four ROIs.

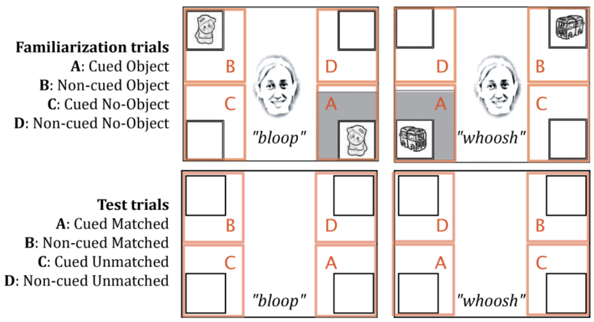

Figure 2: Regions of interest (ROIs) delineated for Familiarization and Test trials.

Results

For each condition (Social Scaffolding and Extended Practice), we first describe infants' looking behavior during the Familiarization trials, followed by the behavior during the Test trials.

A 2 (Cued: Cued or Non-Cued location) x 2 (Location: Object present or No Object) repeated-measures ANOVA[1] was used to examine the effects of cue and audio-visual information on infants' looking distribution during Familiarization. During Test, when infants had to match a sound to now empty locations, the factors of interest were Cue (previously Cued or Non-cued location) and Location (Matched or Unmatched to the sound). A significant Cue x Location interaction for Familiarization trials was taken as evidence that infants followed the cue to objects, and a Cue x Location interaction during Test indicated that infants associated the test sound to only the previously cued matched animation. When a Cue x Location interaction was found, planned t-tests were conducted to confirm that infants looked longer to cued rather than non-cued object present or matched locations.

Social Scaffolding condition (Familiarization trials):

As expected, infants looked predominantly towards the Cued, Object present locations during the Familiarization trials in both the Training and Generalization phases (see Table 1). This was indicated by highly significant Cue x Location interactions (Training: $F(1,15) = 28.43$, $p < .001$, partial η^2=.66; Generalization: $F(1,15) = 24.43$, $p < .001$, partial η^2=.62). Infants looked longer to cued rather than non-cued object-present locations (Training: $t(15) = 5.37$, $p < .001$, Generalization: $t(15) = 5.09$, $p < .001$).

Social Scaffolding condition (Test trials):

Training phase Critically for our hypothesis, we found a significant interaction between Cued and Location during Test, $F(1,15) = 9.01$, $p = .009$, partial η^2=.38. Infants looked longer to cued rather than non-cued matched locations, t(15) = 3.37, p = .004.

Generalization phase Again, as predicted, there was a significant interaction during test between Cued and Location, $F(1,15) = 6.08$, $p = .02$, partial η^2=.29. Infants looked longer to cued rather than non-cued matched locations, t(15) = 2.43, p = .028, which suggests that infants continued to learn from the flashing square in the absence of the scaffolding social signals.

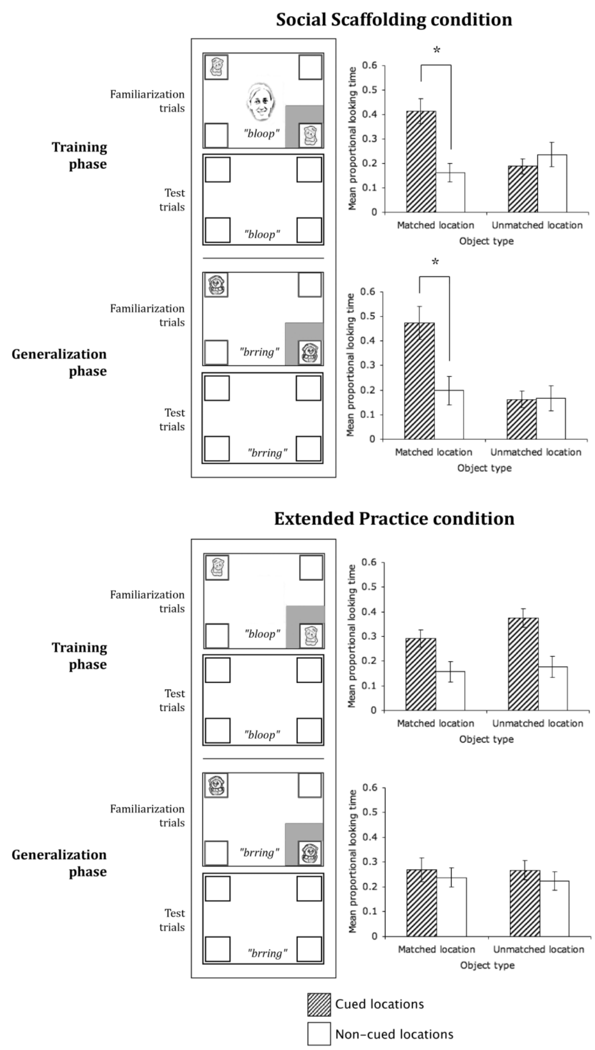

Figure 3. Familiarization and Test trials for Social Scaffolding and Extended Practice conditions. All stimuli were in full color on a black background. *$p < .03$

[1] Though analyzing non-independent looking times violates a basic assumption of ANOVA, including the results from all four locations is the best way for us to conclude that infants show learning within the test trials (e.g., infants look longer in the cued object location than the other locations).

Table 1
Mean proportional looking times during test trials to four regions of interest (ROIs) in both Social Scaffolding and Extended Practice conditions.

	Condition							
	Social Scaffolding condition				Extended Practice condition			
	Training		Generalization		Training		Generalization	
Locations (ROIs)	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Familiarization trials								
Cued object	**0.69**	0.04	**0.65**	0.03	**0.66**	0.03	**0.62**	0.04
Non-cued object	**0.28**	0.04	**0.31**	0.03	**0.30**	0.03	**0.35**	0.03
Cued no-object	**0.02**	0.01	**0.03**	0.01	**0.02**	0.00	**0.02**	0.01
Non-cued no object	**0.00**	0.00	**0.01**	0.00	**0.02**	0.01	**0.01**	0.00
Test trials								
Cued matched	**0.41**	0.05	**0.47**	0.07	**0.29**	0.03	**0.27**	0.05
Non-cued matched	**0.16**	0.04	**0.20**	0.06	**0.16**	0.04	**0.24**	0.04
Cued unmatched	**0.19**	0.03	**0.16**	0.03	**0.37**	0.04	**0.27**	0.04
Non-cued unmatched	**0.24**	0.05	**0.17**	0.05	**0.18**	0.04	**0.22**	0.04

Extended Practice condition (Familiarization trials):

As in the Social Scaffolding condition, infants looked predominantly towards the Cued, Object present location during the Familiarization trials in both the Training and the Generalization phases (see Table 1). There was a highly significant Cue x Location interaction (Training: $F(1,14) = 40.78$, $p < .001$, partial $\eta^2=.74$; Generalization: $F(1,14) = 14.46$, $p = .002$, partial $\eta^2=.51$), and infants looked longer to cued rather than non-cued object-present locations (Training: $t(14) = 5.79$, $p < .001$, Generalization: $t(14) = 3.84$, $p = .002$).

Extended Practice condition (Test trials):

Training phase Contrary to the Social Scaffolding condition, no interaction between Cue and Location was found in the absence of social signals during training ($F(1,14) = 1.07$, $p = .32$, partial $\eta^2=.07$). There was only a main effect of cued location ($F(1,14) = 6.93$, $p = .02$, partial $\eta^2=.33$), where infants looked longer to cued locations than non-cued locations regardless of matching multimodal information ($M_{cued}=.33$, $SE_{cued}=.03$, $M_{non\text{-}cued}=.17$, $SE_{non\text{-}cued}=.03$) (Figure 3).

Generalization phase There were no main effects or interactions during test trials from this phase, suggesting that longer exposure to the novel flashing cue was not responsible for the learning effects observed in the Social Scaffolding condition (Figure 3).

A 2 (Condition: Social Scaffolding or Extended Practice) x 2 (Cued locations) x 2 (Object locations) ANOVA did not reveal a significant interaction for Familiarization trials, $F(1,60) = .44$, $p = .51$, partial $\eta^2 < .01$, showing that the looking distribution during these trials was not affected by the presence of the face in the Social Scaffolding group. Moreover, there was no total looking time difference (in seconds) in the Cued Object present ROI between the Extended Practice condition ($M = 51.82$, $SE = 4.93$) and the Social Scaffolding condition ($M = 49.55$, $SE = 5.17$), $t(29)=.32$, $p=.75$. The ANOVA revealed a significant interaction during Test, $F(1,59) = 12.93$, $p = .001$, partial $\eta^2 =.18$, confirming that infants learned better in the Social Scaffolding condition.

General Discussion

The present study provides the first evidence for a cue-scaffolding process during infancy that is dependent on the presence of social communicative signals. Importantly, this process elicits better learning *about the cued events rather than just better attention to them*. The measure of learning used was infants' anticipatory looks to the cued location of the target event. When the novel flashing cue was paired with a communicative social signal (e.g., an engaging face that maintained mutual gaze), 8-month-olds predicted events would appear in the correct cued locations, even though the face never turned or changed eye gaze. In the absence of these social signals, flashing squares produced only general spatial learning. After initial training that paired flashing squares with communicative signals, infants continued learning from these cues even when the face was no longer present. Critically, this learning effect was not due to more practice with the flashing square cue. Without the initial pairing, infants displayed no learning at this later stage (perhaps losing interest in the cues but being unable to learn about the objects without them, Wu & Kirkham, 2010), suggesting that initial exposure to social signals is necessary to elicit and maintain specific learning from novel attention orienting cues at this age.

In human adults and rats, repeated consistent pairing of cues has been shown to lead to *acquired equivalence*, where the properties of one cue are generalized to its paired cue (Honey & Hall, 1989; Liljeholm & Balleine, 2010). In the early learning environment, social communicative signals (e.g., eye contact or calling someone's name) often precede referential cues (especially eye gaze and pointing; see Corkum & Moore, 1998). Infants could therefore transfer previously acquired knowledge about such social signals (i.e., communicative intent about an upcoming interesting event) to novel referential cues (similar to infants' ability to transfer perceptual grouping rules from one domain to another, Quinn & Bhatt, 2009). In the present study, we propose the infants generalized the knowledge they had about the communicative intent implied by ostensive signals (e.g., mutual gaze, infant-directed speech, smiling face) to the novel attention cues (e.g., flashing squares). Our results provide evidence for a social scaffolding mechanism explaining how infants *learn to learn* from novel cues. This mechanism may help infants eventually learn to use pointing gestures or arrows, cues they do not understand earlier in life.

Other mechanisms have been proposed to explain how infants learn to follow referential cues. For example, it has been proposed that infants can be trained to respond to referential cues through reinforcement learning (e.g., reinforcing appropriate gaze following with an interesting object in the cued location; Corkum & Moore, 1998). Such reinforcement learning is similar to learning the contingencies between cues and upcoming target locations (Johnson, Posner, & Rothbart, 1991; McMurray & Aslin, 2004; Ristic & Kingstone, 2009). The present study, however, goes two steps further, not only showing that infants can learn to use the novel cue in the absence of a more reinforcing target (an equally-salient distracter item was present in all trials), but also showing that cue training supports better learning of the target events.

This mechanism is fast and resistant to distractions, and therefore well adapted to function in the infant's noisy environment. Future studies will have to determine the exact contextual requirements for this scaffolding mechanism to occur, e.g., which of the many social communicative signals that we used are necessary? Acquired equivalence models highlight the role of the consistent repeated pairing of the familiar and the novel cue. Follow-up studies should also measure the extent of the "referential" function of the paired novel cue in a variety of contexts (infants can learn many different types of events from referential cues). We believe the present findings provide an important first step towards elucidating an emerging ability to use various referential cues to support, enhance, and mediate learning, going beyond documenting *which* referential cues guide attention and learning during infancy to proposing a mechanism for *how* this learning occurs.

Acknowledgments

We thank Paul Quinn and Victoria Southgate for helpful comments on earlier versions of this manuscript. Thank you to Leslie Tucker and Marian Greensmith for help with data collection. This research was supported by a grant to RW and NZK from the University of London Central Research Fund and a grant to Mark Johnson from the UK Medical Research Council, G0701484. Presentation of this proceeding was supported by the British Academy Overseas Conference Grant awarded to RW.

References

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology, 34*, 28-38.

Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology, 85*, 199–212.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science, 20*, 347–353.

Gredebäck, G., Melinder, A.M.D., & Daum, M. (in press). The development and neural basis of pointing comprehension. *Social Neuroscience.*

Honey, R. C., & Hall, G. (1989). Acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 338-346.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adults' eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*, 131–134.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience, 3*, 335–344.

Kita, S. (Ed.) (2002). *Pointing: where language, culture, and cognition meet.* Mahwah, NJ: Lawrence Erlbaum.

Leekham, S. R., Solomon, T. L., & Teoh, Y.-S. (2010). Adults' social cues facilitate young children's use of signs and symbols. *Developmental Science, 13*, 108-119.

Liljeholm, M., & Balleine, B. W. (2010). Extracting Functional Equivalence From Reversing Contingencies. *Journal of Experimental Psychology: Animal Behavior Processes, 36*, 165–171.

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant Rule Learning Facilitated by Speech. *Psychological Science, 18*, 387-391.

McMurray, B. and Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6,* 203-229.

Quinn, P. C., & Bhatt, R. S. (2009). Transfer and scaffolding of perceptual grouping occurs across organizing principles in 3- to 7-month-old infants. *Psychological Science, 20,* 933-938.

Ristic, J., & Kingstone, A. (2009). Rethinking attentional development: Reflexive and volitional orienting in children and adults. *Developmental Science, 12*, 289–296.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*, 1–4.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*, 705-722.

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology, 107*, 118-136.

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (in press). Infants learn about objects from statistics and people. *Developmental Psychology*.

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences, 105*, 13690–13695.