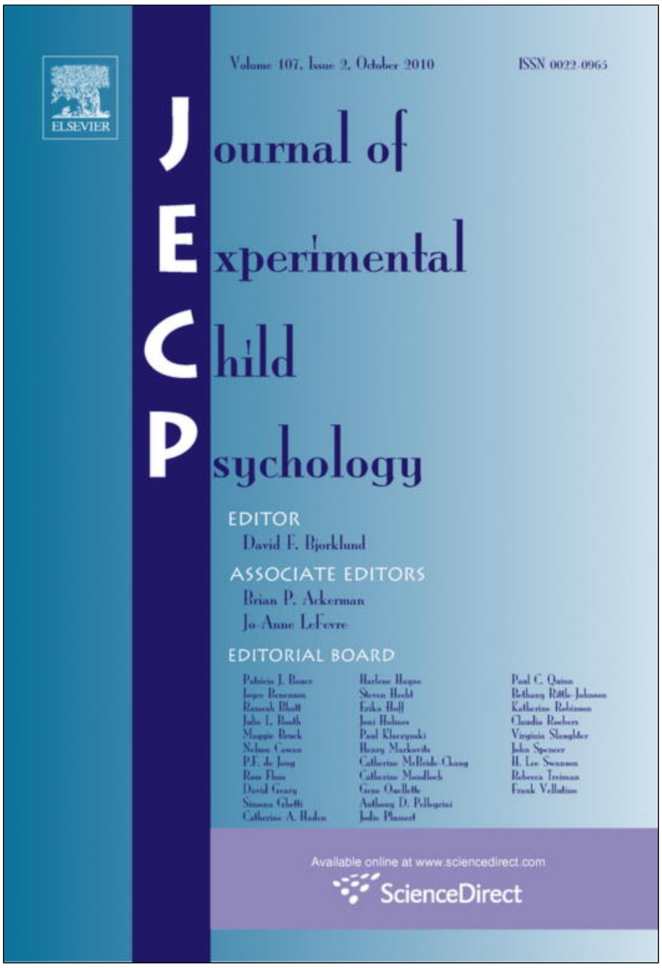
Volume 107, Issue 2, October 2010
ISSN 0022-0965
ELSEVIER
Journal of Experimental Child Psychology
EDITOR
David F. Bjorklund
ASSOCIATE EDITORS
Brian P. Ackerman
Jo-Anne LeFevre
EDITORIAL BOARD
Available online at www.sciencedirect.com
ScienceDirect

Journal of Experimental Child Psychology 107 (2010) 118–136

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

No two cues are alike: Depth of learning during infancy is dependent on what orients attention

Rachel Wu *, Natasha Z. Kirkham

Centre for Brain and Cognitive Development, Birkbeck College, University of London, London WC1E 7HX, UK

ARTICLE INFO

Article history:
Received 4 August 2009
Revised 26 April 2010

Keywords:
Visual attention
Spatial cueing
Infancy
Learning
Multimodal perception
Eye-tracking
Joint visual attention

ABSTRACT

Human infants develop a variety of attentional mechanisms that allow them to extract relevant information from a cluttered multimodal world. We know that both social and nonsocial cues shift infants' attention, but not how these cues differentially affect learning of multimodal events. Experiment 1 used social cues to direct 8- and 4-month-olds' attention to two audiovisual events (i.e., animations of a cat or dog accompanied by particular sounds) while identical distractor events played in another location. Experiment 2 directed 8-month-olds' attention with colorful flashes to the same events. Experiment 3 measured baseline learning without attention cues both with the familiarization and test trials (no cue condition) and with only the test trials (test control condition). The 8-month-olds exposed to social cues showed specific learning of audiovisual events. The 4-month-olds displayed only general spatial learning from social cues, suggesting that specific learning of audiovisual events from social cues may be a function of experience. Infants cued with the colorful flashes looked indiscriminately to both cued locations during test (similar to the 4-month-olds learning from social cues) despite attending for equal duration to the training trials as the 8-month-olds with the social cues. Results from Experiment 3 indicated that the learning effects in Experiments 1 and 2 resulted from exposure to the different cues and multimodal events. We discuss these findings in terms of the perceptual differences and relevance of the cues.

* Corresponding author. Fax: +44 0 20 7631 6587.
E-mail address: r.wu@bbk.ac.uk (R. Wu).

0022-0965/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2010.04.014

Introduction

In a noisy and exciting environment, attention cues help infants to filter the vast amount of possible events worth attending to at a given moment. For example, an infant can follow both the mother's voice ("Look at that, baby!") and pointing finger to a fun toy. Nonsocial cues also shift infants' attention; infants will look toward a flashing siren as they hear and see an ambulance drive by. Because attention is crucial for learning, these types of cues can bias the events that infants learn. Although many cues shift attention, little is known about whether different cues produce different learning effects. Using the principles of central and peripheral cueing from the classic Posner cueing paradigm (e.g., Posner, 1980), we investigated whether central (social) and peripheral (nonsocial) attention-shifting cues elicit differential learning of audiovisual events. Specifically, ensuring equal duration of visual attention (i.e., sustained looking) with both attention cues to the cued target event, this study measured the infants' predictions of the event's location over time.

Attentional cueing

The classic Posner cueing paradigm examines the impact of central and peripheral attention cues on the speed of target detection (Posner, 1980). The paradigm presents participants with a central fixation point and two locations in which targets can appear on the left and right sides of the computer screen. Preceding the target, a cue appears; either a flashing square around a target location (peripheral cue) or an arrow between the locations pointing left or right (central cue) shifts attention to one side. The peripheral cue attracts attention to itself, whereas the central cue directs attention to another location. Across trials, the cues either predict (valid) or do not predict (invalid) the upcoming target's location. A participant's task is to press a key as soon as a target is detected. In this seminal study, the peripheral cues were nonpredictive overall (i.e., equal number of valid and invalid trials), whereas the central cue was predictive (i.e., targets appeared 80% of the time in cued locations) (Posner, 1980). Because participants were faster at detecting targets on valid trials with either cue, it was concluded that peripheral cues elicit reflexive shifts in attention, whereas central cues elicit voluntary shifts.

Several studies have since challenged this original conclusion by demonstrating that nonpredictive central cues (i.e., arrows and eye gaze) do indeed elicit reflexive shifts in attention. Both nonpredictive arrows (Eimer, 1997; Ristic, Friesen, & Kingstone, 2002; Tipples, 2002) and eye gaze cues (Driver et al., 1999; Friesen & Kingstone, 1998; Langton & Bruce, 1999; Schuller & Rossion, 2001) elicit faster target detection in valid trials. Children from 3 to 7 years of age show similar reflexive attention shifts with valid central arrows and eye gaze cues as well as peripheral flashing cues (Akhtar & Enns, 1989; Iarocci & Burack, 2004; Ristic & Kingstone, 2009; Ristic et al., 2002; Senju, Tojo, Dairoku, & Hasegawa, 2004; Varga et al., 2009), although children do not display complete adult-like orienting and performance until approximately 8 years of age (Goldberg, Maurer, & Lewis, 2001; Ristic & Kingstone, 2009). In sum, both central and peripheral cues shift attention efficiently when the cues match the upcoming target's location.

Attentional cueing with infants

The aforementioned studies with children and adults used covert orienting (e.g., shifting visual attention without head or eye movements) and reaction time measures (i.e., button press). In contrast, infancy research tends to use overt orienting (e.g., shifting visual attention with head or eye movements) and looking time measures, which are easier to elicit from infants (Johnson, 1994). Whether infants orient to salient peripheral cues is consistent with age (e.g., a bright light that captures a neonate's attention will also capture that of a 6-month-old), although distracting or sustaining the infant's attention with nonsocial cues depends on age (Cohen, 1972; Colombo, 2001; Richards & Casey, 1992) and characteristics of the stimuli (e.g., stimulus complexity or salience) (Oakes & Tellinghuisen, 1994; Tellinghuisen & Oakes, 1997). Orienting faster to cued locations than to noncued locations with peripheral cues develops quickly from birth and resembles adult-like responses by 4 months of age (Johnson, 1994; Johnson, Posner, & Rothbart, 1991).

Gaze cueing

Central cues in infancy research tend to be gaze cues because infants do not respond to the direction of an arrow (Varga et al., 2009). Eyes and gaze direction are particularly salient for infants. Baron-Cohen (1995) and Baron-Cohen, Campbell, Karmiloff-Smith, Grant, and Walker (1995) underscored this notion by proposing the existence of an eye direction detector (EDD), a neural module that tracks the presence of eyes, eye contact, and gaze direction. Neonates prefer direct rather than averted gaze (Farroni, Csibra, Simion, & Johnson, 2002), whereas the ability to detect gaze direction develops from birth to 4 months of age (Vecera & Johnson, 1995). Studies using an adapted Posner cueing paradigm have found facilitated target detection from gaze direction with a computerized face in 3-month-olds (Hood, Willen, & Driver, 1998) and more recently in newborns (Farroni, Massaccesi, Pividori, Simion, & Johnson, 2004).

Joint attention

Gaze cueing forms the basis of joint attention, roughly defined as ‘‘looking where someone else is looking” (Butterworth, 2004). With a spatial layout similar to the gaze cueing paradigm, the standard joint attention paradigm involves a live person (rather than a computer-animated face) that combines the gaze cue with other social cues such as infant-directed speech, initial eye contact, head turn, and gestures. As a result, the central social cue in these studies is typically more dynamic and perhaps more ecologically valid than the faces used in the gaze cueing paradigm. By 3–4 months of age, infants can follow the ‘‘looker” to the target and do so reliably by 5–6 months of age under simple ideal conditions (Butterworth, 2004; Flom & Pick, 2007; Poulin-Dubois, Demke, & Olineck, 2007; Senju & Csibra, 2008; Striano & Reid, 2006). Interestingly, if eye gaze is isolated from the other cues in this paradigm, only infants older than 18 months of age follow the looker’s gaze direction (Moore & Corkum, 1998), although (as mentioned previously) much younger infants show facilitation effects based on the gaze direction of an animated face (Hood et al., 1998). In sum, compounded live social cues, especially eye gaze and head turn with either infant-directed speech or initial eye contact, become powerful attention mediators during early infancy (Senju & Csibra, 2008). If infants learn to engage in joint attention during the first half year of life, perhaps infants ‘‘learn to learn” from the face as well as they do with other central attention cues during childhood (e.g., arrows) (Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003; Varga et al., 2009).

Given that central/social and peripheral/nonsocial cues shift attention effectively during infancy, how do those shifts affect the depth of processing of an event? If both types of cues can bias infants’ information intake equally, perhaps infants would learn target events from either cue.

Basic learning from social attention cues

A few studies have investigated how social cues guide basic learning in a noisy environment, namely in word and language learning (Akhtar, Carpenter, & Tomasello, 1996; Baldwin, 1993; Carpenter, Nagell, & Tomasello, 1998; Gliga & Csibra, 2009; Goldstein & Schwade, 2008; Houston-Price, Plunkett, & Duffy, 2006; Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006) and object processing (Cleveland, Schug, & Striano, 2007; Gergely, 2008; Striano, Chen, Cleveland, & Bradshaw, 2006; Yoon, Johnson, & Csibra, 2008). Infants may use social cues to solve the problem of referential uncertainty, where one label may refer to multiple objects or events (Houston-Price et al., 2006); this use develops over time from prelinguistic infants to expert word learners during the second year of life (Pruden et al., 2006). By 9 months of age, infants encode different object properties (i.e., identity or spatial location) depending on whether they see a person pointing to or reaching for an object (Yoon et al., 2008). There is neural and behavioral evidence showing that infants during the first year recognize gazed-on objects as more familiar than noncued objects (Cleveland et al., 2007; Reid, Striano, Kaufman, & Johnson, 2004; Striano et al., 2006; Theuring, Gredebäck, & Hauf, 2007). Recently, Wu, Gopnik, Richardson, and Kirkham (2010) found that social cues help infants learn about objects through cued (rather than noncued) feature co-occurrences. These studies suggest that social cues influence learning language and object properties. However, no published study to date has investigated how social and nonsocial attention cues mediate the depth or quality of learning in a noisy multimodal environment.

Multimodal learning

Audiovisual events abound in infants' surroundings (e.g., seeing and hearing people talking, toys falling, cars driving). The ability to understand that some sights and sounds belong together allows infants to organize their cluttered world (Bahrick, 2004; Lewkowicz, 2000). One important advantage of having veridical multimodal representations is that they can help to predict future events. For example, when information that had been presented in two modalities (e.g., seeing and hearing someone walking down the stairs) is now presented in only one modality (e.g., hearing someone walking down the stairs), it is useful for infants to know that somebody will appear at the bottom of the stairs momentarily (Bahrick, 2004).

Richardson and Kirkham (2004, Experiment 2) measured infants' ability to predict the appearance of a visual event after familiarizing infants to the audiovisual event and then providing only the audio information (dynamic spatial indexing). Familiarization trials presented infants with two white frames on the left and right sides on a black background. Each of the two target events appeared in its own frame without temporal overlap. These events displayed objects moving synchronously to particular sounds. For example, on one familiarization trial, an infant could see a toy cat twisting to a "bloop" sound in the left frame. On the following trial, the infant could see a toy dog moving to a "boing" sound in the right frame. These events were repeated while the sounds, objects, and locations maintained their redundant mappings. After six repetitions of the familiarization trials, the test trials displayed the two blank frames while one of the two sounds played in the absence of the visual stimuli. For a given test trial, one square was matched to the sound (i.e., where an object had appeared previously moving to the same sound), whereas the other square was the unmatched location. The 6-month-olds in this experiment looked longer to the matched square than to the unmatched square, indicating that the infants bound the audiovisual events. Recently, this finding was extended to 3-month-olds (Kirkham, Richardson, Wu, & Johnson, 2010).

This paradigm displayed only one audiovisual event at a time (i.e., without distractions presented simultaneously). In the current study, we adapted the dynamic spatial indexing paradigm by adding a distractor that was identical to the target event in every familiarization trial. In so doing, we presented infants with two possible locations to bind the sound. In addition, infants were presented with either a central/social cue, a peripheral/nonsocial cue (directing attention to one of the two identical events), or no cue. We intended for the cues to shift infants' attention to the target event for the same duration. By presenting two valid locations for audiovisual binding within a trial and two cued locations across all trials, we could distinguish cueing effects from multimodal learning as well as the interaction of the two factors. The four possible learning outcomes were learning only from cues (i.e., anticipating events in previously cued locations regardless of event type), learning from the audiovisual events (i.e., anticipating events in both valid binding locations), a combined outcome (i.e., looking to cued target locations), and neither (i.e., looking equally to all possible locations).

This study investigated whether social and nonsocial cues (similarly effective but characteristically different attention cues) would enhance multimodal learning of the cued target event in an ambiguous noisy environment. Because both cues can shift infants' attention effectively to an event in a particular location, perhaps these different ways of increasing the target's salience can produce similar depths of learning. The following three experiments investigated how these cues impact infants' predictions of upcoming audiovisual events. To assess how social cues mediate learning at different ages, Experiment 1 (*face cue condition*) directed 8- and 4-month-olds' attention to two audiovisual events (i.e., animations of a cat or dog accompanied by particular sounds) while identical distractor events played in another location. We included these age groups in this condition because infants start engaging in joint visual attention by 3 or 4 months of age and do so reliably by 8 months of age (e.g., Butterworth, 2004). To assess how nonsocial cues mediate learning, Experiment 2 used flashing squares wrapped around the target event to draw 8-month-olds' attention to that location. The first condition in Experiment 2 (*one-color cue condition*) displayed a red flashing square wrapped around the target frame for both audiovisual events (i.e., cat and dog). Because infants are drawn to the salience of these peripheral cues when shifting attention (Oakes & Tellinghuisen, 1994), perhaps their intense salience could encourage infants to bind the cue to the sounds rather than the objects to the sounds. Thus, a second condition (*two-color cue condition*) assigned a different color cue to each event type (e.g., red for cat trials and yellow for dog trials) to highlight the differences in the events. Providing infants with the additional color cue may elicit the specific learning effects that the one-color cue may hinder. Among the 8-month-old conditions with an attention cue, we controlled for the duration of looking to the cued target event during familiarization so that gross attention or inattention (e.g., due to cue salience) would not be a factor influencing learning. Experiment 3 measured baseline responses to learning without attention cues either through exposure to both familiarization and test trials (*no cue condition*) or through exposure to only the test trials (*test control condition*).

Considering the effectiveness of both peripheral/nonsocial and central/social cues, we predicted that the 8-month-olds in the face cue, one-color cue, and two-color cue conditions would look longer at the target event than at the distractor during familiarization. For the test trials, we predicted that these infants would anticipate the object's appearance in its cued target location rather than in the other three locations. We predicted that the 4-month-olds in the face cue condition (with less experience in learning from social cues) would follow the face during familiarization but later exhibit poor learning effects. Finally, we predicted that without attention cues, infants would look equally to the locations containing objects during familiarization and then look longer to those same locations when hearing the appropriate sound (correct object locations) than to the locations that were matched to the other object (incorrect object locations). We expected equal looking to all four frames without familiarization trials in the test control condition.

Experiment 1

Experiment 1 assessed the usefulness of social cues at different ages in directing attention and learning in a competing salience situation. In this experiment, 8- and 4-month-olds were eye-tracked in the face cue condition, which showed a face that turned to one of two identical multimodal events (Fig. 1, top panel). The simultaneously presented multimodal events served as the noisy environment within which infants needed to choose where to focus their attention.

Methods

Participants

In total, 29 full-term 8-month-olds (mean age = 8 months 6 days, range = 7 months 19 days to 8 months 28 days, 12 girls and 17 boys) and 23 full-term 4-month-olds (mean age = 4 months 5 days, range = 3 months 23 days to 4 months 24 days, 11 girls and 12 boys) participated in this experiment (face cue condition). A further 4 infants were excluded from final analyses due to calibration errors ($n = 3$) or fussiness (i.e., failing to complete more than one block, $n = 1$). The majority of infants ($n = 43$) were of White middle socioeconomic status. The infants were recruited via magazine advertisements and flyers and were given T-shirts or bibs as "thank you" gifts for their participation.

Apparatus

Infants' looks were monitored using a Tobii 1750 eye-tracker. All dynamic stimuli were presented on the 17-inch monitor attached to the Tobii eye-tracking unit using Tobii's ClearView AVI presentation software with sounds played through stereo external speakers (for technical details about the eye-tracker, refer to von Hofsten, Dahlström, & Fredriksson, 2005). The experimenter monitored whether infants were attending to the screen through an external video camera mounted on top of the Tobii screen. Infants' looks were recorded with the ClearView software. The animated object clips were created using Adobe Photoshop 7 and Macromedia Director MX 2004 (Richardson & Kirkham, 2004), and the live face clip was filmed using Macintosh iMovie (version 4.0.1). All movie clips were assembled using Final Cut Express HD 3 (Apple).

Stimuli

Multimodal events. Following the paradigm from Richardson and Kirkham (2004), we trained infants on two types of events bound to two different sounds. This study modified the original paradigm by simultaneously presenting a target and an identical distractor event on every trial (instead of just the

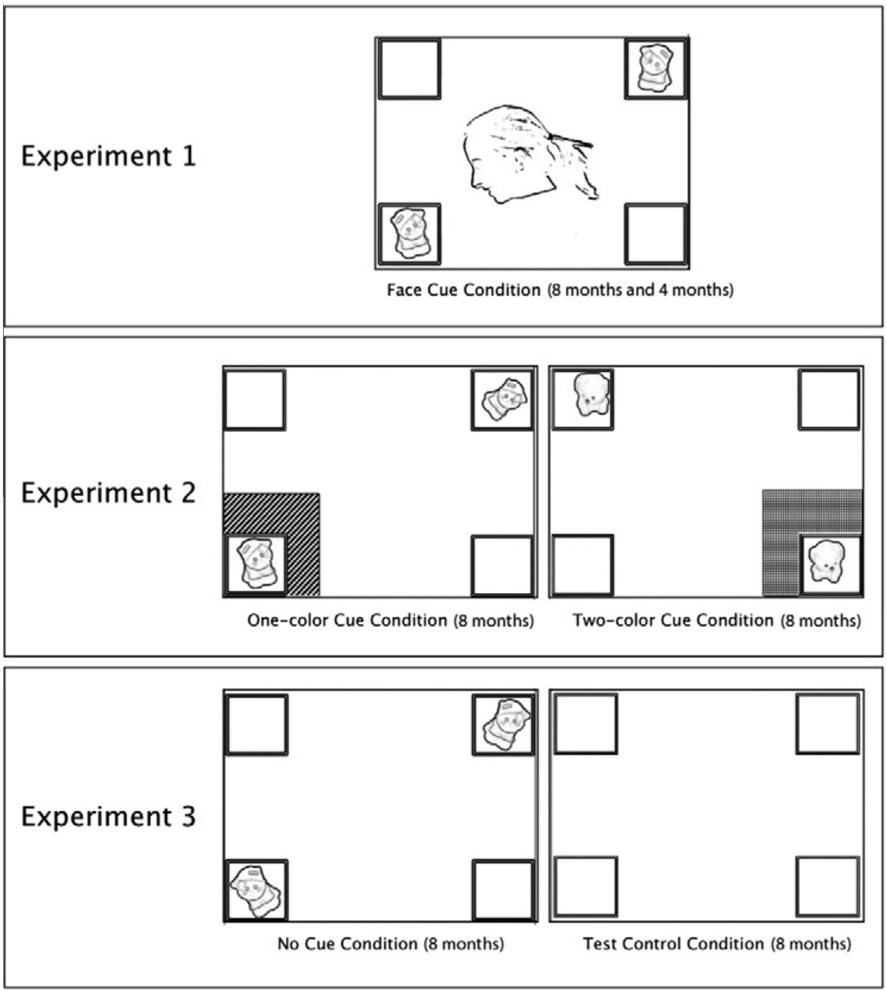

Fig. 1. Sample familiarization events in all five conditions. The diagonal stripes represent a red flashing square, and the hatched lines represent a yellow flashing square in the middle panel (Experiment 2). All stimuli were in full color.

target event). Infants watched two pairs of audiovisual events (i.e., two orange toy cats or two blue toy dogs) moving in synchrony to two sounds during familiarization (i.e., the cats moved to a "bloop" tune, and the dogs moved to a "boing" sound). The target and distractor were identical to equate their visual salience, although their features and movements were mirrored to help infants perceive the events as separate. Each pair was presented separately; infants either saw two cats or two dogs on the screen during a given familiarization trial in diagonally opposite corners. During one familiarization trial, for example, one infant would see two identical cats appear in the lower left and upper right corners dancing to the bloop tune. On another trial, the infant would see two identical dogs appear in the lower right and upper left corners moving to the boing sound. With four overall locations across trials (one in each corner of the screen), these object events did not overlap spatially. These events were contained within white frames on a black background and lasted for 6 s.

Social cue. The central social attention cue in the face cue condition was a film of a live actress. The actress was a young Caucasian woman with her hair pulled back in a ponytail, and she spoke with a local accent. Because she wore a black shirt and was filmed against a black background, the film only showed her head and neck. The female face looked out at the baby with a neutral expression, smiled, said "Hi baby, look at this!" in infant-directed speech, and then turned to look down at either the lower left or lower right corner. This portion of the clip lasted for 5 s with no other visual stimulus on the screen except for the four blank white frames in each corner. The woman's turned smiling face remained on the screen while the multimodal events appeared and played for 6 s. After the multimodal events disappeared from the screen, the smiling face immediately turned back to the center and changed back to a neutral expression.

Previous cueing studies have typically used two target locations (left and right in the middle of the screen, or left and right at the bottom of the screen). Our actress turned to the lower left and right corners to maintain consistency with the methods used in recent joint visual attention studies (e.g., Cleveland et al., 2007; Gergely, 2008; Gergely, Egyed, & Királyi, 2007; Senju & Csibra, 2008; Striano et al., 2006). This paradigm required the use of four locations (uncommon in the cueing literature). The actress cued the infant to the lower two corners instead of the middle of the screen to maximize the spatial segregation of the target and distractor multimodal events and to avoid spatial overlap between trials.

Design and procedure

After a short session in a lab waiting room during which infants were acclimated to the experimenter and the lab setting, infants were tested individually while sitting on the caregiver's lap in a small quiet room. Infants were seated 50 cm away from the monitor. Prior to starting the experiment, the experimenter used a five-point calibration on infants' looks. Infants were shown a Bert and Ernie *Sesame Street* clip in the four corners and center of the screen until they fixated in each location. Infants who did not calibrate on at least four points were excluded from the final analysis ($n = 2$) (for technical details about the calibration procedure, refer to von Hofsten et al., 2005).

The familiarization trials in this experiment (face cue condition) had two main phases. First, the face cue appeared on the screen with the four blank frames, spoke to the infant, and turned; this portion of the clip lasted for 5 s. Then the multimodal events appeared immediately as the face finished turning and played for 6 s in diagonally opposite corners while the face stayed in the turned position. The face always cued the infant to a valid target. If the face turned to the lower right corner, objects appeared in the lower right and upper left corners; if the face turned to the lower left corner, objects appeared in the lower left and upper right corners. After the multimodal events played for 6 s (with the face still turned on the screen), they disappeared as the face turned back to look at the infant again. At that point, the trial ended, and an attention-getter was displayed until the infant's fixation returned to the screen for at least 1500 ms. The test trials presented only the four blank white frames (without objects) on a black background and played either the bloop or boing sound for 5 s.

Infants were presented with four stimuli blocks (six familiarization trials and one test trial per block). The six familiarization trials displayed three events per object–sound pairing along with the central stimulus (face). The test trial played one of the sounds to measure saccades to the associated object's cued location. Although the order of the familiarization trials was identical across all four blocks, the test trials played only one of the two sounds (one sound in Block 1 and the other sound in Block 2). Blocks 3 and 4 repeated Blocks 1 and 2. Over four blocks, infants viewed 24 familiarization trials separated by four test trials. The trials were counterbalanced between infants for presentation of object location and test sound. Presentation of object order was pseudo-randomized. No event was repeated more than two times in succession, so that the test sound was heard in either the last or second-to-last familiarization trial (e.g., ABABBA, BAABAB) (Richardson & Kirkham, 2004). Centrally presented attention-getters (i.e., still kaleidoscopic circle or square with different ring sounds) followed every trial presentation. The experimenter turned it off to play the next trial when the infant had looked at the attention-getter for approximately 1500 ms (see Fig. 2 for a schematic of the experimental design).

Coding

The total looking times to the four corners and central area of the screen during familiarization and test trials were recorded with the eye-tracker and compiled by Tobii's ClearView analysis software. Fixations that were shorter than 100 ms were excluded from the final analyses. Proportional looking time to the areas of interest (AOIs) in each trial for every infant was calculated by dividing the total looking time to that location by the total looking time to every location. Proportional looking time rather than total looking time was used to avoid measuring learning effects that were driven by infants who looked longer overall than the others.

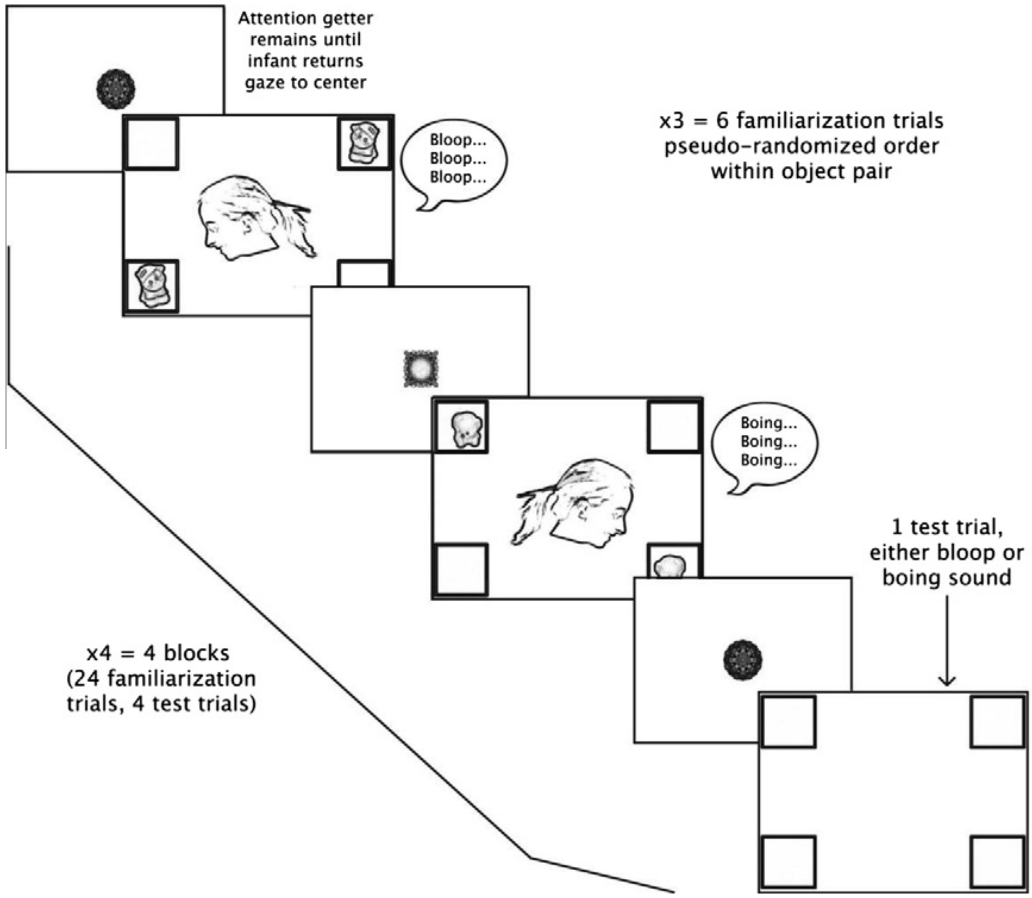

Fig. 2. Schematic of the experimental design for all conditions in Experiments 1 through 3 except for the test control condition. On the computer screen, the background was black, the frames around the objects were white, and the objects and attention cues were in full color.

The four corners of the screen were labeled for familiarization and test trials based on whether the location had objects or were cued during familiarization; two corners displayed objects that were associated with a particular sound (correct object locations) (*object*), and two corners were cued locations across all familiarization trials (*cued*) (see Fig. 3). Using two valid locations for binding per trial (target and distractor with one sound) and cueing infants to two different locations during all familiarization trials allowed us to distinguish cueing effects (i.e., looking to cued locations regardless of audiovisual binding) from multimodal learning effects (i.e., correct audiovisual binding) during test trials. To examine learning over time during the test trials (*block*), we split the experimental session into two halves (Blocks 1–2 and Blocks 3–4) because the second half repeated the first half.

The center of the screen was also recorded as an AOI because it impacted overall proportional looking time in the face cue condition. Because the face cue was present during the entire familiarization trial, infants sometimes looked at it while the objects were on the screen. The looking time to the center portion of the screen was included in calculating the proportional looking time but was not included in the final analyses because we were interested in the fixations in the corners of the screen.

Results

Analyses were conducted to quantify the effectiveness of the social cue in directing infants' attention to and learning about the target. For a given familiarization trial, there were two correct locations in which infants could look: to the frames containing the identical cats and the identical dogs. Within a block of familiarization trials, infants were cued to two locations (one for the cat trials and one for the dog trials). Therefore, for the familiarization trials, a 2 (Object: correct or incorrect) $\times$ 2 (Cued: cued or noncued) repeated-measures analysis of variance (ANOVA) was conducted over the mean proportional looking times during all four blocks (i.e., four blocks of 6 familiarization trials totaling

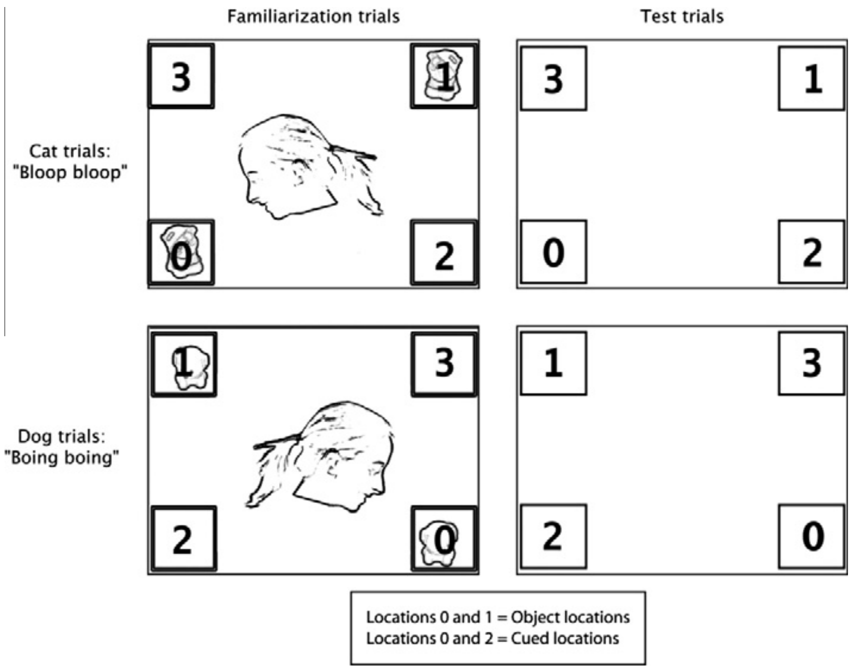

Fig. 3. Correct object and cued locations for sample familiarization and test trials. Using two valid locations for binding per trial (target and distractor with one sound) and cueing infants to two different locations during all familiarization trials allowed us to distinguish cueing effects (i.e., looking to cued locations regardless of audiovisual binding) from multimodal learning effects (i.e., correct audiovisual binding).

24 familiarization trials). Any interactions revealed by the ANOVA for the looking times during familiarization trials were further investigated through Bonferroni-corrected pairwise comparisons between the cued and noncued correct object locations because we were only interested in how infants attended to the locations with visual stimuli during familiarization.

To investigate learning effects (based on cueing, audiovisual binding, or both) over time during the test trials, an omnibus 2 (Block: Blocks 1–2 or Blocks 3–4) $\times$ 2 (Object: correct or incorrect) $\times$ 2 (Cued: cued or noncued) repeated-measures ANOVA was conducted over the mean proportional looking times during all test trials. If the ANOVA revealed a two-way interaction involving block, we split the data by the first two and last two blocks and used Bonferroni-corrected pairwise comparisons to investigate differences in looking to appropriate and inappropriate locations (cued/noncued or object correct/incorrect). For this analysis, the data were collapsed (averaged) accordingly based on the second factor (cued or object). If the ANOVA revealed a two-way interaction involving cued and object, we collapsed the data across all blocks and used Bonferroni-corrected pairwise comparisons to compare the mean proportional looking time in the cued versus noncued locations for both the correct and incorrect object locations. If the ANOVA revealed a three-way interaction, we split the data by the first two and last two blocks and used the same Bonferroni-corrected pairwise comparisons to compare the mean proportional looking time in the cued versus noncued locations for both the correct and incorrect object locations. Because we were interested in analyses that would reveal subtle differences among looking times in four locations with each cue, we only investigated behavior within conditions rather than across conditions.

Face cue condition (8 months)

The 2 $\times$ 2 repeated-measures ANOVA on the familiarization trials revealed main effects of object and cued location, with significantly longer looking to both correct object locations (rather than incorrect object locations) and cued locations (rather than noncued locations), $F_{\text{object}}(1, 28) = 5.61\text{E}2$, $p_{\text{object}} < .001$, partial $\eta^2 = .95$, $F_{\text{cued}}(1, 28) = 12.76$, $p_{\text{cued}} = .001$, partial $\eta^2 = .31$.[1] There was an interaction

[1] The F values for the familiarization trials are large because the vast majority of infants looked only to the areas with objects and followed the face. Therefore, the between participant variability was tiny.

between object and cued locations, $F_{object \times cued}(1, 28) = 16.12$, $p < .001$, partial $\eta^2 = .37$. To investigate the interaction, we ran Bonferroni-corrected pairwise comparisons. Infants looked longer to cued target locations (the object the face turned to) than to noncued object locations during familiarization ($p = .001$) (see Fig. 4). Table 1 lists the mean proportional looking times to the four corners and center of the screen during the familiarization trials.

A 2 × 2 × 2 ANOVA investigating learning over blocks during test trials revealed a marginal main effect of cued location, $F_{cued}(1, 28) = 3.87$, $p = .059$, partial $\eta^2 = .12$, an interaction between block and cued location, $F_{block \times cued}(1, 28) = 4.42$, $p = .045$, partial $\eta^2 = .14$, and an interaction for block, cued, and object location, $F_{block \times cued \times object}(1, 28) = 8.54$, $p = .007$, partial $\eta^2 = .23$. We split the data from the first two and last two blocks to investigate the two-way interaction involving the cued factor. The collapsed (averaged) proportional looking times during the first two blocks to the cued and noncued locations showed that infants did not look significantly longer to either cued or noncued locations ($p = .42$, $M_{cued} = .20$, $SE_{cued} = .02$, $M_{noncued} = .17$, $SE_{noncued} = .02$). In Blocks 3 and 4, infants looked longer to cued locations than to noncued locations ($p < .01$, $M_{cued} = .22$, $SE_{cued} = .02$, $M_{noncued} = .12$, $SE_{noncued} = .02$).

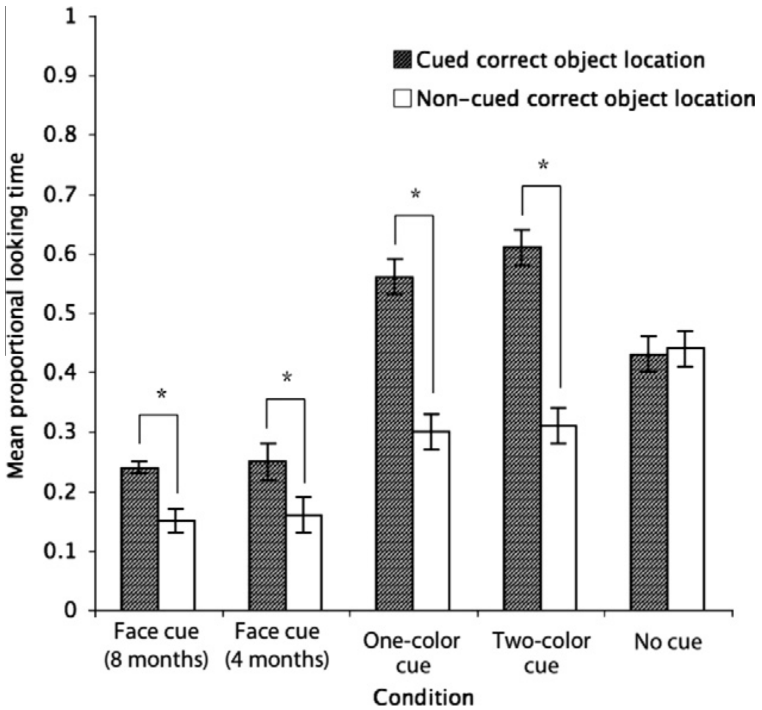

Fig. 4. Mean proportional looking times to the cued correct object location and noncued correct object location during familiarization trials from Blocks 1 through 4 in the four conditions. $^*p \leqslant .05$.

Table 1
Mean proportional looking times during familiarization trials from Blocks 1 through 4 to five areas of interest across four conditions: face cue (8 and 4 months), one-color cue, two-color cue, and no cue.

Locations (AOIs)	Condition									
	Face cue				One-color cue		Two-color cue		No cue	
	8 months		4 months		8 months		8 months		8 months	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Familiarization trials										
Cued, correct object	0.24	0.01	0.25	0.03	0.56	0.03	0.61	0.03	0.43	0.03
Noncued, correct object	0.15	0.02	0.16	0.03	0.30	0.03	0.31	0.03	0.44	0.03
Cued, incorrect object	0.02	0.00	0.02	0.00	0.04	0.00	0.01	0.00	0.03	0.00
Noncued, incorrect object	0.01	0.00	0.02	0.01	0.04	0.01	0.01	0.00	0.04	0.01
Center	0.58	0.02	0.54	0.03	0.06	0.00	0.06	0.00	0.06	0.00

Critically for our hypothesis, we ran Bonferroni-corrected comparisons to investigate the three-way interaction. We compared the mean proportional looking time in the cued versus noncued locations for both the correct and incorrect object locations. During the first two blocks, infants did not look more to cued locations than to noncued locations for both correct ($p = 1.00$) and incorrect ($p = .46$) object locations. By Blocks 3 and 4, infants looked more to cued locations than to noncued locations for the correct object locations ($p = .01$) but not for the incorrect object locations ($p = 1.00$). Fig. 5 graphs these mean proportional looking times, and Table 2 displays the looking times to the AOIs from the test trials in the first two and last two blocks.

Face cue condition (4 months)

Results from a 2 × 2 repeated-measures ANOVA on the familiarization trials revealed a significant main effect for object location, $F_{object}(1, 22) = 3.65E2$, $p_{object} < .001$, partial $\eta^2 = .94$, a marginal effect for cued location, $F_{cued}(1, 22) = 3.71$, $p_{cued} < .07$, partial $\eta^2 = .14$, and an interaction between the two, $F_{object \times cued}(1, 22) = 4.93$, $p_{object \times cued} = .04$, partial $\eta^2 = .18$. Infants looked longer in locations with objects (rather than no objects) and showed a trend for looking longer to cued locations (rather than noncued locations). Post hoc comparisons investigating the interaction revealed that infants looked longer to cued object locations than to noncued object locations during familiarization ($p = .05$) (see Fig. 4).

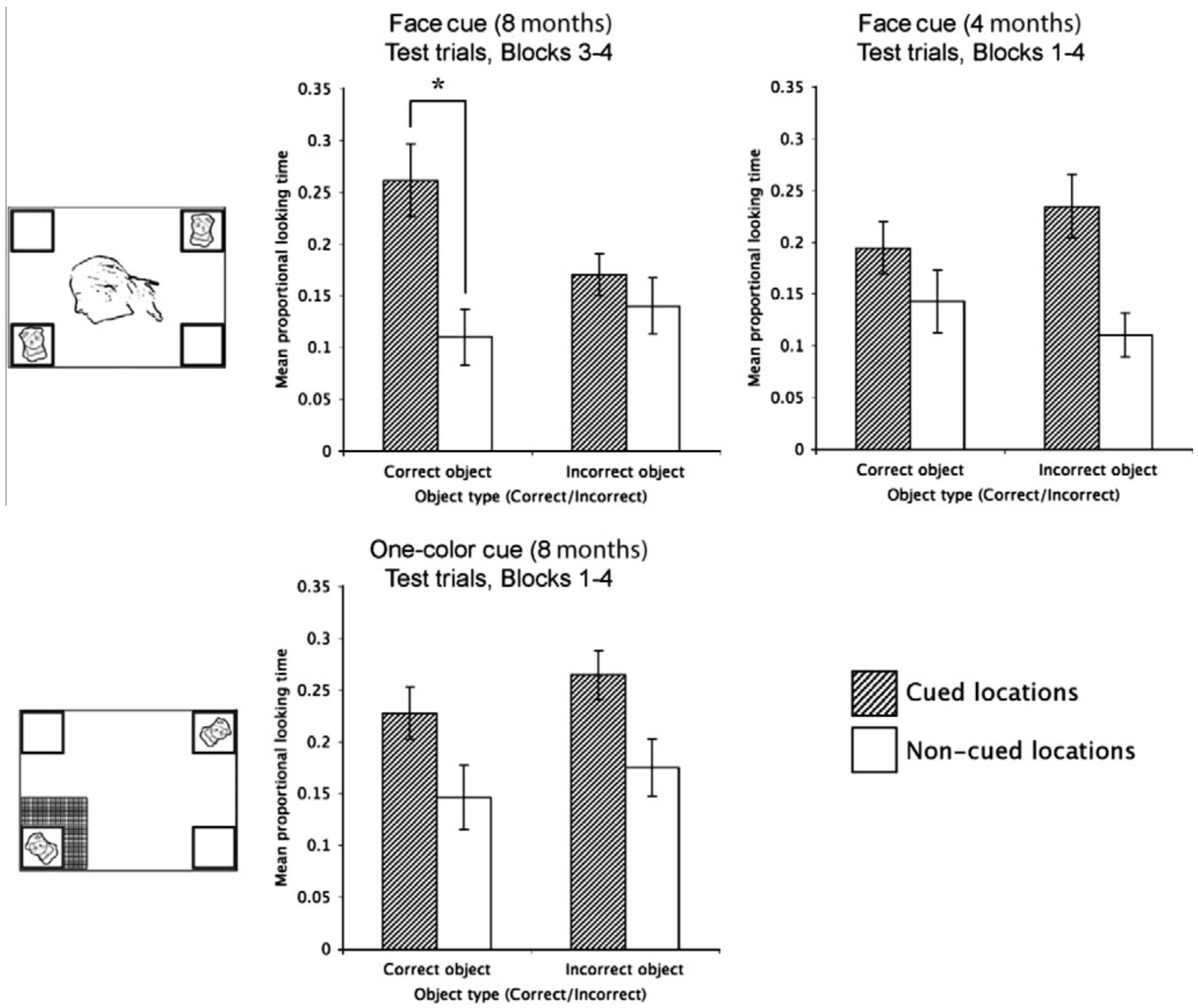

Fig. 5. Mean proportional looking times to locations during test trials from Blocks 3 and 4 in the face cue condition (8 months) and Blocks 1 to 4 in the face cue condition (4 months) and one-color cue condition (8 months). The 8-month-olds in the face cue condition looked more to the cued correct object location, whereas the 4-month-olds in the same condition looked longer to cued locations than to noncued locations regardless of object–sound mappings. The 8-month-olds in the one-color cue condition (flashing squares) also looked longer only to cued locations than to noncued locations regardless of multimodal information. *$p = .01$.

Table 2
Mean proportional looking times during test trials from Blocks 1 and 2 and Blocks 3 and 4 to five areas of interest across four conditions: face cue (8 and 4 months), one-color cue, two-color cue, and no cue.

Locations (AOIs)	Condition									
	Face cue				One-color cue		Two-color cue		No cue	
	8 months		4 months		8 months		8 months		8 months	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Test trials from Blocks 1 and 2										
Cued, correct object	0.15	0.03	0.25	0.04	0.22	0.04	0.23	0.04	0.18	0.03
Noncued, correct object	0.18	0.03	0.16	0.04	0.14	0.03	0.17	0.03	0.18	0.03
Cued, incorrect object	0.25	0.03	0.24	0.04	0.30	0.04	0.21	0.04	0.21	0.03
Noncued, incorrect object	0.16	0.03	0.11	0.03	0.17	0.03	0.18	0.04	0.23	0.03
Center	0.25	0.03	0.23	0.03	0.18	0.04	0.20	0.03	0.20	0.03
Test trials from Blocks 3 and 4										
Cued, correct object	0.26	0.04	0.14	0.03	0.24	0.03	0.23	0.04	0.23	0.04
Noncued, correct object	0.11	0.02	0.13	0.04	0.16	0.03	0.12	0.03	0.18	0.03
Cued, incorrect object	0.17	0.03	0.22	0.04	0.23	0.03	0.22	0.03	0.12	0.03
Noncued, incorrect object	0.14	0.03	0.10	0.02	0.18	0.03	0.19	0.03	0.21	0.04
Center	0.32	0.04	0.41	0.06	0.19	0.03	0.24	0.03	0.25	0.05

A $2 \times 2 \times 2$ repeated-measures ANOVA across all test trials revealed two significant main effects for block and cued locations. For the main effect of block, F_{block} (1, 22) = 10.71, p = .003, partial η^2 = .33, infants looked proportionally less to the four frames overall during the second half of the study (Blocks 3–4) than during the first half of the study (Blocks 1–2) because more time was spent looking at the central AOI (where the face had been during familiarization), $M_{\text{blocks1-2}}$ = .77, $SE_{\text{blocks1-2}}$ = .03, $M_{\text{blocks3-4}}$ = .59, $SE_{\text{blocks3-4}}$ = .06. For the main effect of cued location, F_{cued} (1, 22) = 5.31, p = .03, partial η^2 = .19, infants looked longer to cued locations than to noncued locations regardless of audiovisual mappings across all blocks, M_{cued} = .43, SE_{cued} = .05, M_{noncued} = .25, SE_{noncued} = .04. Because this analysis revealed no two-way or three-way interactions (involving block, cued, and/or object), no other analyses were carried out.

Discussion

As predicted, both 8- and 4-month-olds followed the face cue in the intended direction during familiarization. However, only the 8-month-olds reliably predicted that the objects would appear in cued locations associated with the correct object when they heard the appropriate sound during test trials. The 4-month-olds showed general spatial learning by looking to both cued locations regardless of the sound played. These results suggest that the social cue is an efficient attention director by 4 months of age and perhaps is an effective learning tool by 8 months of age, helping infants to recognize, recall, and predict appropriate events. The 8-month-olds' use of this cue required time to develop over an experimental session because the infants as a group only learned the appropriate locations by the third and fourth blocks. The 4-month-olds may have exhibited a poorer learning effect because they did not use the social cue in the same way as their older peers or because they could not process the complex array as efficiently as the 8-month-olds. Perhaps with more familiarization trials or a simplified array, these younger infants could show specific learning effects similar to those of the 8-month-olds. A follow-up study with either modification would be necessary to identify the appropriate conclusion. These two conditions highlight the developed potential of the social cue to help infants process their noisy environment.

Experiment 2

In Experiment 1, a central face cue directed 8-month-olds' attention effectively to the intended target and produced appropriate audiovisual binding of the cued event. Because infants show facilitated target detection with both valid central cues (e.g., eye gaze) and peripheral cues (e.g., flashing squares), perhaps the latter can produce the same learning effect as faces. Unlike the success of a central cue, the success of a peripheral cue in shifting attention depends on its salience because it attracts attention to itself. Over two conditions, Experiment 2 cued infants with bright flashing squares presented around the target frame (Fig. 1, middle panel). The first condition (one-color cue condition) presented infants with a bright red flashing square wrapped around the target frame while the target and distractor audiovisual events played. This condition measured the impact of a salient peripheral cue on multimodal binding. However, the cue's intense salience could have encouraged infants to bind the cue to the sounds rather than the objects to the sounds. In other words, the salient visual cue could have overridden the featural differences in the objects, similar to labels overriding featural categorization (Plunkett, personal communication; Plunkett, Hu, & Cohen, 2008). Therefore, we included a second condition (two-color cue condition) that presented each event with its own colored square (red or yellow) to help the infants further differentiate the two events.

Methods

The apparatus, stimuli, design, and procedure for Experiment 2 were identical to those for Experiment 1 except for the attention cue used. The central face stimulus was replaced with a peripheral flashing square wrapped around the target frame.

Participants

In total, 60 full-term 8-month-olds participated in this experiment (mean age = 8 months 4 days, range = 7 months 15 days to 8 months 27 days): one-color cue condition ($n = 30$, 14 girls and 16 boys) and two-color cue condition ($n = 30$, 13 girls and 17 boys). The majority of infants ($n = 48$) were of White middle socioeconomic status.

Stimuli

The multimodal events with the cats and dogs in this experiment were identical to those in Experiment 1. The cue in these trials was a peripherally presented flashing square wrapped around the target frame in either the lower left or lower right corner (locations that the face had cued in Experiment 1). As in the face cue condition, the cue was always valid, with the onset of the flashing square occurring simultaneously with the multimodal event. Each familiarization trial was 6 s long. The cue in the one-color cue condition was a red square for both multimodal events, whereas the two-color cue condition used a red or yellow square, one for the cued cat and one for the cued dog. As in Experiment 1, object location and test sound (along with the color of the cue in the two-color cue condition) were counterbalanced between participants.

Results

As in Experiment 1, proportional looking times to all corners and the central area of the screen were calculated, and corner looks were used in the final analyses.

One-color cue condition

During the one-color familiarization trials, an ANOVA revealed main effects of object, F_{object} (1, 29) = 1.90E3, $p_{\text{object}} < .001$, partial $\eta^2 = .99$, and cued location, F_{cued} (1, 29) = 14.78, $p_{\text{cued}} = .001$, partial $\eta^2 = .34$. Infants looked longer to locations with objects (rather than no objects) and to those that were cued (rather that noncued). There was also an interaction between the two, $F_{\text{object} \times \text{cued}}$ (1, 29) = 28.00, $p < .001$, partial $\eta^2 = .49$. A Bonferroni-corrected pairwise comparison revealed that the infants looked longer to cued object locations than to noncued object locations ($p < .001$). For this condition, a separate analysis was carried out to determine whether infants were looking only to the red flashing square instead of to the target object during familiarization trials. If they looked long enough to the object, the flashing square did not compete significantly for gross visual attention. When drawn to cued locations during familiarization, the infants spent the majority (83.65%) of their time looking to the actual object rather than to the red flashing square. A one-way ANOVA revealed that total looking time (in seconds) to the target objects in this condition did not differ from that in the face cue condition, $F(1, 57) = .28$, $p = .60$, partial $\eta^2 = .01$, $M_{face} = 54.76$, $SE_{face} = 4.08$, $M_{one\text{-}color} = 51.68$, $SE_{one\text{-}color} = 4.20$.

A repeated-measures ANOVA across all test trials revealed a significant main effect only of cued location, F_{cued} (1, 29) = 4.08, $p = .05$, partial $\eta^2 = .12$. Infants looked longer to cued locations than to noncued locations regardless of the audiovisual mappings across all blocks. Because this ANOVA revealed no two-way or three-way interactions (involving block, cued, and/or object), no other analyses were carried out.

Two-color cue condition

During familiarization for the two-color cue condition, there were main effects of object, F_{object} (1, 29) = 5.28E3, $p_{object} < .001$, partial $\eta^2 = 1.00$, and cued location, F_{cued} (1, 29) = 22.32, $p_{cued} < .001$, partial $\eta^2 = .44$, as well as an interaction between the two, $F_{object \times cued}$ (1, 29) = 27.48, $p_{object \times cued} < .001$, partial $\eta^2 = .49$. Infants looked longer to locations with objects (rather than no objects) and to cued locations (rather than noncued locations). A Bonferroni-corrected pairwise comparison revealed that the infants looked longer to cued object locations than to noncued object locations ($p < .001$). During the test trials, there were no main effects or interactions.

Discussion

As predicted, during familiarization, the infants followed both the one-color and two-color flashing square cues to the intended location. However, depth of processing in these conditions was poor, and only general spatial learning occurred in the test trials over the four blocks in the one-color cue condition. During the test trials, the infants in this condition looked at both of the cued locations regardless of the sound played, resembling the learning effect from the face cue condition with the 4-month-olds. These results suggest that even though the peripheral cue is efficient in directing attention (not just to itself but also to objects), it might not produce the same depth of learning as the central social cue. Importantly, infants looked equally long to the object (as opposed to the flashing square) in this condition as in the previous face cue condition (Experiment 1). Therefore, learning was not compromised by reduced looking to the cued target object. The partial learning effect did not match our original prediction of obtaining full learning effects. Even the two-color flashing squares (i.e., extra redundant cues) did not help the infants to differentiate the events. One explanation for this finding is that perhaps the different colors offered yet another feature to bind to the object–sound event and, thus, overloaded the infants. In addition, the asynchrony between the flashing frequency and tempo of the bloop and boing sounds may have also overloaded the infants.

Experiment 3

Experiment 3 measured baseline learning without attention cues over two conditions: no cue condition (with familiarization and test trials) and test control condition (with only test trials) (Fig. 1, bottom panel). The no cue condition displayed only the multimodal events during familiarization and then the standard test trials. To ensure that the familiarization trials affected the looking behaviors during the test trials from the other conditions, the test control condition presented infants with only the test trials displaying the four blank frames that played either the bloop or boing sound.

Methods

Participants

In total, 44 full-term 8-month-olds participated in this experiment (mean age = 8 months 2 days, range = 7 months 15 days to 9 months 1 day): no cue condition ($n = 28$, 13 girls and 15 boys) and test control condition ($n = 16$, 6 girls and 10 boys). The majority of infants ($n = 32$) were of White middle socioeconomic status.

Stimuli

The stimuli in the no cue condition were identical to those in the one-color and two-color cue conditions except that there was no attention cue; the identical multimodal events were presented for 6 s per trial between attention-getters. The test trials for the no cue condition were identical to those for every other condition. For the no cue condition, cued target locations were labeled as the lower corners with an object (previously cued locations in Experiments 1 and 2). The test control condition showed infants only the blank test trials with sound (two blocks of consecutive test trials alternating between the bloop and boing sounds) to determine looking preferences in the absence of any visual stimuli.

Results

No cue condition

During familiarization with no attention cue, there was only a main effect of object on looking time as predicted, F_{object} (1, 27) = 5.54E3, $p < .001$, partial $\eta^2 = 1.00$, with longer looking to object locations than to nonobject locations. During the test trials in all four blocks, a $2 \times 2 \times 2$ ANOVA revealed an interaction between block and object locations, $F_{\text{block} \times \text{object}}$ (1, 27) = 4.07, $p = .05$, partial $\eta^2 = .13$. To investigate this interaction, we used pairwise comparisons to investigate differences in looking times to the locations that previously contained or did not contain the correct objects (appropriate or inappropriate multimodal mappings) between the first two and last two blocks. During the first two blocks, infants looked longer to incorrect locations than to correct locations, although this difference was not significant ($p = .17$, $M_{\text{object}} = .18$, $SE_{\text{object}} = .02$, $M_{\text{no-object}} = .22$, $SE_{\text{no-object}} = .03$). In Blocks 3 and 4, infants looked longer to correct object locations than to incorrect object locations, although again this difference was not significant ($p = .22$, $M_{\text{object}} = .21$, $SE_{\text{object}} = .02$, $M_{\text{no-object}} = .17$, $SE_{\text{no-object}} = .02$). The two-way interaction from the ANOVA was due to a shift from looking in incorrect object locations to looking in correct object locations.

Test control

For the test control condition where infants saw only the blank frames and heard either the cat or dog sound, a repeated-measures ANOVA revealed that the mean looking times to the four corners of the screen had no significant differences, $F(1, 16) = 0.51$, $p = .68$, partial $\eta^2 = .03$, $M_{\text{right-down}} = .16$, $SE_{\text{right-down}} = .03$, $M_{\text{right-up}} = .19$, $SE_{\text{right-up}} = .04$, $M_{\text{left-down}} = .25$, $SE_{\text{left-down}} = .05$, $M_{\text{left-up}} = .22$, $SE_{\text{left-up}} = .04$.

Discussion

Without an attention cue during familiarization (no cue condition), infants looked equally long to both objects. During the test trials, infants in the no cue condition gradually shifted from looking in incorrect to correct object locations. However, as a group, they failed to look significantly longer to correct locations by the last two blocks. Although infants as young as 3 months of age can recall the locations correctly when shown one pairing at a time (Kirkham et al., 2010), doubling the visual array and valid binding locations creates a more complex scene to process and learn. Therefore, it is not too surprising that even 8-month-olds had trouble with this complex audiovisual scene if they were not provided with an attention cue. Random looks in the test control condition confirmed that any looking time differences during the test trials from the previous four conditions were due to variations in the familiarization trials.

General discussion

This study investigated how infants learn from different attention cues. The experiments implemented a paradigm based on central and peripheral cueing concepts to determine whether two effective attention cues during infancy help learning in a distraction-filled environment. We found that central/social and peripheral/nonsocial attention cues impact infants' ability to bind basic multimodal events. Interestingly, infants showed reliably different learning effects at test across conditions despite attending for a similar duration to the target event with either cue. The 8-month-olds following the social cue during familiarization anticipated that objects would appear during test in the cued correct object locations, where faces had attended to correct objects during familiarization. Although the 4-month-olds also followed the face cue, they anticipated events only in the cued locations regardless of the event type. These younger infants remembered only where (rather than what) they were cued to during familiarization. The flashing square cues (one-color or two-color) shifted 8-month-olds' attention admirably but elicited poor learning effects. Similar to the 4-month-olds with the social cue, the one-color cue produced no more than a shallow anticipation of events in cued locations. Even the red and yellow squares (two-color cue condition), which were intended to help infants further distinguish the multimodal events, appeared to suppress significant learning effects. Without a cue, 8-month-olds as a group had difficulty with the audiovisual binding: They showed only a slight shift toward anticipating events in both valid locations.

This study is the first to show that learning can be dependent on the nature and presence of attention cues as well as on the age of the infants. Our findings with the social cue extend a growing literature showing that following social cues encourages the learning of cued events by 8 or 9 months of age (e.g., Goldstein & Schwade, 2008; Wu et al., 2010; Yoon et al., 2008). Besides demonstrating that 8-month-olds recalled the specific events that the face had attended, we also found a developmental trajectory where 4-month-olds recalled only where the face had attended. Therefore, we concluded that perhaps infants "learn to learn" from the social cue over the first year of life based on studies demonstrating facilitated target detection from gaze cues within the first few months (faster orienting to cued locations [Hood et al., 1998]) and studies on the gradual emergence of sustained joint visual attention by the end of the first year (looking to a gazed-on object [Butterworth, 2004; Carpenter et al., 1998]). One argument against this conclusion from our findings is that the younger infants may have had difficulty in processing the complex display (3-month-olds are successful with the simpler original paradigm) (Kirkham et al., 2010). If we had implemented a simplified array or provided more familiarization trials, perhaps these younger infants would have displayed more sophisticated anticipations. With 8-month-olds, peripheral/nonsocial cues elicited learning of the cues' locations rather than properties of the target event. Our findings suggest that the information that infants extract from a scene depends on the infants' age and type of cue present. Using paradigms with various cues and multiple age groups would allow investigations into the usefulness of different cues for infants in different situations.

One issue we did not fully address in this study (because it was not part of our initial aim) is the confounds in the cues. This issue is especially pertinent given the different learning effects that we obtained with each cue. Although these issues are beyond the scope of this article, we offer a few speculations. In this study, we used the central and peripheral cues that have been successful in shifting infants' attention in previous studies with similar paradigms. However, these cues are fundamentally different. First, the cues are a combination of factors that do not overlap between cue types; the only trait they share is that they both shift infants' attention with similar efficiency. Our central cue is social, and the peripheral cue is nonsocial, as is the case in most of the infant cueing studies. Unless a future study disentangles these elements, one cannot determine whether the learning effects were due to the location of the cue, their social/nonsocial nature, or both. Moreover, the social cue is an audiovisual compound cue. It is composed of many other cues, namely infant-directed speech, eye contact, head turn, expression of interest, and gaze. Each of these sub-cues is a viable learning tool, especially speech (Fulkerson & Waxman, 2007; Xu, 2002). Perhaps one sub-cue can elicit the same learning effects as the compounded cue, and altogether maybe these cues are more useful than a less compounded peripheral cue. Second, the familiarity of the cue may have induced the different learning effects. Perhaps 8-month-olds learned from the social cue because they have had more experience with that cue than have 4-month-olds. Moreover, perhaps 8-month-olds did not learn as well from the flashing square cue as they did from the social cue because flashing squares were not familiar to them. To address whether cue familiarity plays a role in learning, a future study could familiarize infants to learn from flashing squares. Finally, the physical and temporal distances between the cue and the target are different for each cue type. This difference is especially important because learning in this study required binding temporally synchronous multimodal events in particular locations. Because peripheral cues attract (rather than direct) attention, perhaps they competed temporally and spatially for attentional resources. In our paradigm, the flashing square conditions had a simultaneous cue–target onset and the square consistently flashed around the target to maintain infants' gaze on the multimodal event. A pilot condition using a cue onset prior to the target onset as well as having the flashing stop at the target onset did not maintain infants' attention to the target multimodal event as the face had done and, therefore, was discontinued. Because of the spatial and temporal competition induced by the flashing squares on the target event, infants may have had more difficulty in binding the appropriate audiovisual events. That said, infants did look comparably long to the multimodal events with both types of cues. Therefore, the qualitative difference in learning was not a function of looking time to the objects during familiarization. Perhaps our central face cue, which directed attention from itself onto other events, provided the spatial and temporal distance necessary to process the intended target. Although these perceptual differences of the cues highlight the need to disentangle the confounds, our finding still stands: Different cues, similarly effective in orienting attention, elicit different learning outcomes.

As a final speculation, let us put aside the fundamental perceptual differences in the two cues and assess their instructive roles and their relevance to learning about other events. Perhaps flashing cues are more often used to draw attention to themselves so that the observer learns about the cue presence rather than about the properties of an event it surrounds. For example, if a flashing siren appears in view, one learns quickly to steer away from it rather than learn about the lettering painted on the ambulance. However, faces have a lifelong role of instructing children about their external world. Some even argue that social cues are essential for infant learning. For example, Kuhl (2007) proposed that social factors "gate" infant language acquisition; that is, social interactions direct how humans learn specific communicative skills. If the different types of cues are meant to elicit different types of learning, one cannot conclude that a particular learning effect is poorer or less sophisticated than the other.

By studying how infants learn events from attention cues, we investigated the usefulness of attentional biases for infants' cognitive development. Cues can reside at the level of the target (audiovisual redundancy in this study) and of an external attention mediator (e.g., faces, flashing squares). Both levels provide solutions to recognizing and learning appropriate events in ambiguous situations. Understanding the interactions of cues at different levels will provide a more accurate picture of infants' learning mechanisms in different situations (e.g., dynamic systems theory [Thelen & Smith, 1994], emergentist coalition model [Hollich, Hirsh-Pasek, & Golinkoff, 2000]). For instance, word learning can be grounded through real-time social interactions, where certain perceptual and motor patterns involving the social partner and objects in the environment predict successful word mappings (Yu, Smith, & Pereira, 2008). Gray areas still exist regarding developmental learning mechanisms, but many answers lie in investigating infants' use of cues at multiple levels in a rich environment.

Acknowledgments

We gratefully acknowledge the efforts of the infants and parents who participated in this study. We also thank Teodora Gliga, Daniel Richardson, Mark Johnson, Denis Mareschal, Eddy Davelaar, Chen Yu, Daniel Yurovsky, and Celeste Kidd for helpful discussions on this work as well as Regina Law, Sara Serbin, and Amanda Weiner for their help with coding. Last, we thank Jodie Plumert and two anonymous reviewers for their insightful comments during the review process. This research was supported in part by a grant to N.Z.K. from the British Academy (SG-47879) and a grant to Mark Johnson from the U.K. Medical Research Council (G0701484).

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development, 67*, 635–645.

Akhtar, N., & Enns, J. T. (1989). Relations between covert orienting and filtering in the development of visual attention. *Journal of Experimental Child Psychology, 48*, 315–334.

Bahrick, L. E. (2004). Development of intermodal perception. In L. Nadel (Ed.). *Encyclopedia of cognitive science* (Vol. 2, pp. 614–617). London: Nature Publishing Group.

Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*, 395–418.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology, 13*, 379–398.

Butterworth, G. (2004). Joint visual attention in infancy. In A. Slater & G. Bremner (Eds.), *Theories of infant development* (pp. 317–354). Oxford, UK: Wiley–Blackwell.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4, Serial No. 255).

Cleveland, A., Schug, M., & Striano, T. (2007). Joint attention and object learning in 5- and 7-month-old infants. *Infant and Child Development, 16*, 295–306.

Cohen, L. B. (1972). A two process model of infant visual attention. In *Paper presented at the Merrill–Palmer conference on research and teaching of infant development*. Detroit, Michigan, USA.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337–367.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*, 509–540.

Eimer, M. (1997). Uninformative symbolic cues may bias visual-spatial attention: Behavioral and electrophysiological evidence. *Biological Psychology, 46*, 67–71.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 9602–9605.

Farroni, T., Massaccesi, S., Pividori, D., Simion, F., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5*, 39–60.

Flom, R., & Pick, A. D. (2007). Increasing specificity and the development of joint visual attention. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (pp. 95–112). Mahwah, NJ: Lawrence Erlbaum.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*, 490–495.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105*, 218–228.

Gergely, G. (2008). Learning "about" versus learning "from" other minds: Natural pedagogy and its implications. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind. Foundations and the future* (Vol. 3, pp. 170–198). New York: Oxford University Press.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science, 10*, 139–146.

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

Goldberg, M., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: The effects of endogenous cueing and of distractors. *Developmental Science, 4*, 209–219.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*, 515–523.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3, Serial No. 262)).

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

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology, 95*, 27–55.

Iarocci, G., & Burack, J. A. (2004). Intact covert orienting to peripheral cues among children with autism. *Journal of Autism and Developmental Disorders, 34*, 257–264.

Johnson, M. H. (1994). Visual attention and the control of eye movements in early infancy. *Attention & Performance, 15*, 291–310.

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.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. (2003). Attention researchers: It's time to look at the real world! *Current Directions in Psychological Science, 12*, 176–184.

Kirkham, N. Z., Richardson, D., Wu, R., & Johnson, S. P. (2010). The importance of "what": Infants use featural information to index events. Birkbeck College, University of London, submitted for publication.

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Developmental Science, 10*, 110–120.

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition, 6*, 541–568.

Lewkowicz, D. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126*, 281–308.

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology, 16*, 495–503.

Oakes, L. M., & Tellinghuisen, D. J. (1994). Examining in infancy: Does it reflect active processing? *Developmental Psychology, 30*, 748–756.

Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*, 665–681.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Poulin-Dubois, D., Demke, T. L., & Olineck, K. M. (2007). The inquisitive eye: Infants' implicit understanding that looking leads to knowing. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (pp. 263–282). Mahwah, NJ: Lawrence Erlbaum.

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development, 77*, 266–280.

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *NeuroReport, 15*, 2553–2555.

Richards, J. E., & Casey, B. J. (1992). Development of sustained visual attention in the human infant. In B. A. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 30–59). Mahwah, NJ: Lawrence Erlbaum.

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General, 133*, 46–62.

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review, 9*, 507–513.

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

Schuller, A.-M., & Rossion, B. (2001). Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Cognitive Neuroscience, 12*, 2381–2386.

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

Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry, 45*, 445–458.

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology, 3*, 289–299.

Striano, T., & Reid, V. M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences, 10*, 471–476.

Tellinghuisen, D. J., & Oakes, L. M. (1997). Distractibility in infancy: The effects of distractor characteristics and type of attention. *Journal of Experimental Child Psychology, 64*, 232–254.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: Bradford Books/MIT Press.

Theuring, C., Gredebäck, G., & Hauf, P. (2007). Object processing during a joint gaze following task. *European Journal of Developmental Psychology, 4*, 65–79.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review, 9*, 314–318.

Varga, K., Frick, J. E., Stansky, L., Beck, H., Dengler, M. J., & Bright, M. A. (2009). I saw the sign: The development of directional understanding. In *Poster presented at the meeting of the society for research in child development*, Denver, CO.

Vecera, S. P., & Johnson, M. H. (1995). Eye gaze detection and cortical processing of face: Evidence from infants and adults. *Visual Cognition, 2*, 59–87.

von Hofsten, C., Dahlström, E., & Fredriksson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. *Infancy, 8*, 217–231.

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2010). Social cues support learning about objects from statistics in infancy. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the cognitive science society* (pp. TBD). Austin, TX: Cognitive Science Society.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*, 223–250.

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.

Yu, C., Smith, L., & Pereira, A. (2008). Grounding word learning in multimodal sensorimotor interaction. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), Proceedings of the 30th annual conference of the cognitive science society (pp. 1017–1023). Washington, DC: Cognitive Science Society.