PAPER

WILEY **Developmental Science**

Top-down contextual knowledge guides visual attention in infancy

Kristen Tummeltshammer | Dima Amso

Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, RI, USA

Correspondence
Kristen Tummeltshammer, Department of Cognitive, Linguistic, & Psychological Sciences, Brown University, Box 1821, Providence, RI 02912, USA.
Email: kristen_tummeltshammer@brown.edu

Abstract

The visual context in which an object or face resides can provide useful top-down information for guiding attention orienting, object recognition, and visual search. Although infants have demonstrated sensitivity to covariation in spatial arrays, it is presently unclear whether they can use rapidly acquired contextual knowledge to guide attention during visual search. In this eye-tracking experiment, 6- and 10-month-old infants searched for a target face hidden among colorful distracter shapes. Targets appeared in *Old* or *New* visual contexts, depending on whether the visual search arrays (defined by the spatial configuration, shape and color of component items in the search display) were repeated or newly generated throughout the experiment. Targets in *Old* contexts appeared in the same location within the same configuration, such that context covaried with target location. Both 6- and 10-month-olds successfully distinguished between *Old* and *New* contexts, exhibiting faster search times, fewer looks at distracters, and more anticipation of targets when contexts repeated. This initial demonstration of contextual cueing effects in infants indicates that they can use top-down information to facilitate orienting during memory-guided visual search.

RESEARCH HIGHLIGHTS

- Six- and 10-month-old infants orient attention using rapidly acquired top-down knowledge about the structure of the visual environment.
- Further, top-down contextual knowledge facilitates infants' visual search behavior, as evidenced by shorter latencies, less looking at distracters, and more target anticipation.
- The presence of contextual cueing effects in young infants demonstrates the successful coordination of rudimentary learning, memory, and attention skills: namely, the rapid extraction of patterns of spatial covariation, maintenance and retrieval of task-relevant information in working memory, and performance of a simple search for a physically salient target.

1 | INTRODUCTION

The ability to prioritize attention to relevant spatial locations is critical for meaningful engagement with, and survival in, a rapidly changing multisensory world. As objects move in and out of view and

distractions loom, visual selective attention may determine whether important events are seen and encoded or missed completely. For example, viewers may fail to detect highly salient events, like a person in a gorilla suit running through a basketball game, if visual attention is sufficiently occupied by a competing demand, such as counting the number of passes (Simons & Chabris, 1999). This may in turn affect which information enters the system for subsequent learning and memory. Top-down knowledge about the contexts, rules, goals, and semantics that govern behavior can be used to increase the efficiency of attention orienting when multiple stimuli compete or are in conflict. Human adults and some non-human animals have demonstrated sensitivity to visual contextual cues, drawing on their knowledge of the visual environment to guide orienting, reduce distraction, and facilitate search (Brockmole & Henderson, 2006; Chun & Jiang, 1998, 1999; Goujon & Fagot, 2013; Wasserman, Teng, & Brooks, 2014).

Consider an infant seeking comfort from a parent at a crowded holiday dinner table; arduously scanning a sea of shifting faces, all with relatively similar surface features, and becoming increasingly distressed as her parent's familiar smile bustles around the table. By orienting to her parent's usual seating location or the kitchen door from

where her parent often appears, the infant substantially increases her chances of success, while minimizing effort and search time. This deceptively simple behavior involves learning the structure of a given environment, retrieving and maintaining that information in working memory, and deploying visual attention efficiently against distraction to locate and identify the target. Given the complexity of the processing demands involved, it is not surprising that studies with children have produced mixed results as to whether the ability to use top-down visual contextual knowledge develops gradually with executive control of attention (Dixon, Zelazo, & De Rosa, 2010; Merrill, Conners, Roskos, Klinger, & Klinger, 2013; Vaidya, Huger, Howard, & Howard, 2007; Yoshida, Darby, & Burling, 2011). The present study investigates the early emergence of visual contextual cueing, considering whether benefits of a stable context on visual attention may be apparent already in the first year of life.

The rich covariational structure that exists between visual objects and their global contexts has been shown to facilitate visual processing in adults: for example, faster detection and identification of component objects in coherent, semantically related contexts (Biederman, Mezzanotte, & Rabinowitz, 1982; Davenport & Potter, 2004; Palmer, 1975), and faster search for targets presented in previewed scenes (Goujon, 2011). In the original contextual cueing paradigm developed by Chun and Jiang (1998), adults were presented with visual search displays in which the configurations of distracters predicted the embedded target's location. Participants found the targets faster when configurations of distracters were repeated, indicating that the contextual information facilitated their search for the target. Although infants as young as 3 months have demonstrated sensitivity to the covariational structure of multi-element arrays, it is presently unclear whether they can *use* this rapidly acquired knowledge in service of more efficient visual search.

This is because successful use of top-down contextual knowledge requires the coordinated efforts of learning, memory, and attention systems that are still developing in infancy. First, infants must be able to rapidly learn the stable covariational structure of a given visual environment. A number of studies have demonstrated that young infants are indeed sensitive to stable covariational and predictive structures in visual input (e.g., Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). Three-month-old infants use spatiotemporal regularities to facilitate attention to upcoming visual events, and by 8 months, they show evidence of orienting to predictive, rather than non-predictive, spatial cues (Kirkham, Slemmer, Richardson, & Johnson, 2007; Tummeltshammer & Kirkham, 2013; Tummeltshammer, Mareschal, & Kirkham, 2014; Wentworth, Haith, & Hood, 2002). Recent work has demonstrated that older infants are sensitive to covariation between a target location and the spatial configuration of non-target elements, looking longer when targets appeared in familiar locations compared to novel locations (Bertels, San Anton, Gebuis, & Destrebecqz, 2017).

Second, infants must successfully retrieve and maintain information about environmental structure in working memory. From 4 months of age, visual short-term memory (VSTM) tasks show that infants can plan, execute and correct eye movements by rapidly encoding and maintaining visual information across brief temporal and spatial gaps. Although initially limited to as little as one item at a time, VSTM undergoes rapid development between 6 and 12 months as infants demonstrate the ability to store information about the locations and identities of multiple items (Kaldy & Leslie, 2005; Oakes, Hurley, Ross-Sheehy, & Luck, 2011; Pelphrey et al., 2004; Reznick, Morrow, Goldman, & Snyder, 2004). For example, Pelphrey and colleagues (2004) observed a substantial increase in correct visual search for a hidden object in a multi-element array between 8 and 12 months. Gains in memory encoding have also been noted with developmental increases in orienting efficiency or with the support of external cues (Markant & Amso, 2013; Ross-Sheehy, Oakes, & Luck, 2011).

Third, infants must deploy visual selective attention among multiple competitors in a manner that maximizes search efficiency and minimizes distraction. That is, acquiring the covariational structure and retrieving it from memory does not guarantee being able to efficiently use it for action. Indeed, what infants know is not always reflected in their behavior, as Munakata (1998) demonstrated on the classic A-not-B task, where infants represent a hidden object's correct location but still act according to some previously relevant feature. The attention orienting literature shows that prior to 4 months, infants' attention is directed by simple perceptually salient cues that facilitate orienting (Amso & Johnson, 2006; Johnson, Posner, & Rothbart, 1991; Johnson & Tucker, 1996; Markant & Amso, 2016) and enable success on pop-out search (i.e., rapid detection of a unique target among homogeneous distracters; Adler & Orprecio, 2006; Amso & Johnson, 2006; Colombo, Ryther, Frick, & Gifford, 1995). Between 4 and 6 months, improvements in oculomotor control are evidenced by more efficient scanning, attention switching, and disengagement; in particular, infants exhibit orienting to a stimulus while suppressing attention to competing distraction, a necessary skill for more effortful visual search (Amso & Johnson, 2008; Butcher, Kalverboer, & Geuze, 2000; Johnson & Tucker, 1996).

Motivated by this evidence from separate learning, memory, and attention orienting literatures, we hypothesize that by 6 months, infants have some of the rudimentary skills needed to extract and use contextual regularities in guiding visual search. We also expect that performance will improve between 6 and 10 months, since many of the tasks that demonstrate these skills in 6-month-olds also show substantial development in the second half of the first year (e.g., Amso & Johnson, 2008; Kirkham et al., 2002; Kirkham et al., 2007; Oakes et al., 2011; Pelphrey et al., 2004; Reznick et al., 2004). In an eye-tracking adaptation of the Chun and Jiang (1998) visual search paradigm, we presented 6- and 10-month-old infants with search displays in which they had to detect and orient to an engaging target (i.e., a face) among multiple distracters. Targets appeared in *Old* or *New* visual contexts, depending on whether the visual search arrays (defined by the spatial configuration, shape and color of component items in the search display) were repeated or newly generated throughout the experiment. Targets in *Old* contexts appeared in the same location within the same configuration, such that the context covaried with the target's location. We tested whether infants would discriminate between *Old* and *New* contexts, extract the relevant cues, and use this top-down knowledge to guide their visual search.
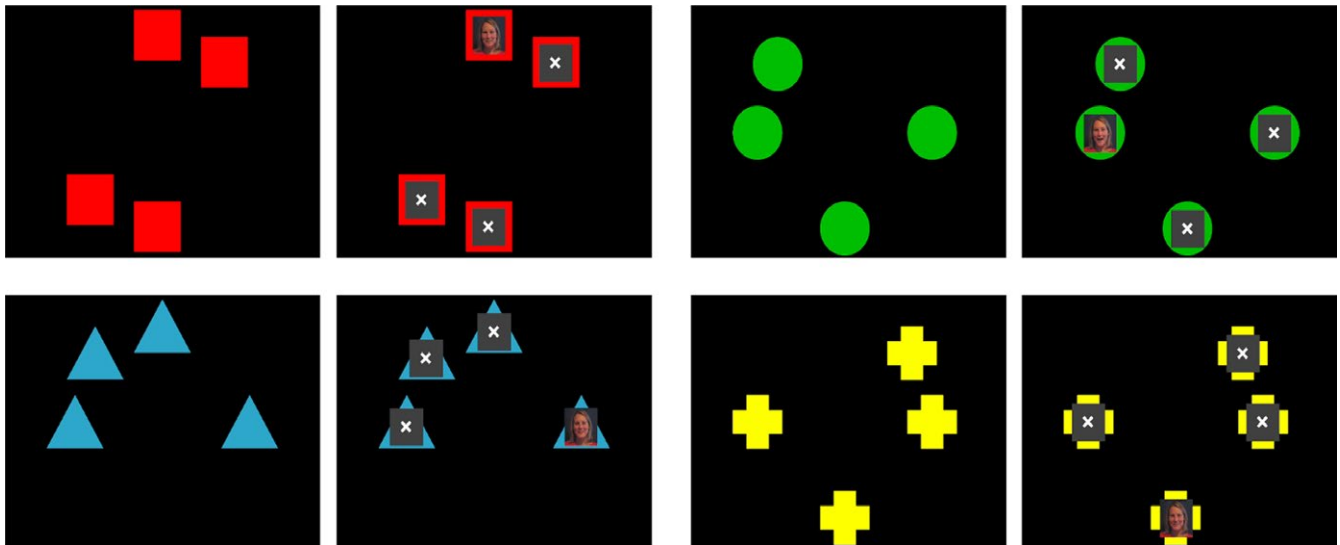
FIGURE 1 Four sample displays of colored shapes, which appeared to flip over and reveal search arrays of a target face and three distracters

If infants are sensitive to the covariation between targets and their contexts and can apply this top-down knowledge to inform their visual search, then we would expect to see faster orienting to targets (i.e., shorter latencies) and fewer looks to distracters presented in *Old* compared to *New* contexts. Given that 6-month-old infants have demonstrated sensitivity to covariation, but have limited ability to maintain and use this information for memory-guided visual orienting and search, we expected contextual cueing effects to be weaker in 6-month-olds compared to 10-month-olds.

2 | METHOD

2.1 | Participants

Forty-six healthy full-term infants participated in the experiment: 25 6-month-olds (11 females, *M* = 199.8 days, *SD* = 13.8 days) and 21 10-month-olds (10 females, *M* = 302.5 days, *SD* = 14.7 days). According to parental report, 33 participants were Non-Hispanic White, five were Hispanic White, three were Black, two were Asian, and three did not report a race/ethnicity. Sample size was determined based on previous studies using the same age groups in our lab and data collection continued until this target sample size was reached. Three additional 6-month-olds and two additional 10-month-olds were tested but not included, due to inattention and/or poor calibration. Infants were recruited via local advertisements and birth records, and informed consent was received from all caregivers. Families received compensation for their time and travel.

2.2 | Apparatus and stimuli

Eye movements were recorded using a remote eye tracker (SensoMotoric Instruments RED system) with a 22″ monitor. Stimuli were presented using the SMI Experiment Center software at a

resolution of 1920 × 1080 pixels, and sounds were played through external stereo speakers. A digital video camera with infrared night vision (Canon ZR960) was placed above the monitor to observe and record infants' head movements.

Infants were presented with search displays of colored shapes: red squares, blue triangles, green circles, or yellow crosses. Each search display contained four items, which could appear in any of eight locations equidistant from the center of the screen (see Figure 1). On each trial, the shapes appeared to flip over, revealing a target stimulus and three distracters. The target stimulus was a smiling female face, which was static for 2 seconds and then became dynamic, addressing the infant with one of three phrases ("Hi baby!", "Great job!", or "Peekaboo!"). The distracter stimuli were gray rectangles with white X's in the center, sized to the same dimensions as the target stimuli (see Figure 1). These stimuli were chosen to ensure that the targets would be highly salient for infants compared to the distracters, prompting infants to search for them. The stimuli were filmed, edited, and animated using Adobe Flash and Premiere Pro software packages.

2.3 | Design and procedure

The two main variables were context (Old vs. New) and Trial (1–12). There were two Old contexts, each consisting of a particular configuration that repeated 12 times throughout the entire experiment. The target stimulus always appeared in the same location within the configuration. There were also two New contexts, which consisted of different configurations that were newly generated on each trial to serve as a control baseline. To rule out any location probability effects, the target appeared equally often in each of four possible locations throughout the experiment (top, bottom, left, right): two locations were used in the Old contexts, and the other two were used in the New contexts. Hence, any difference in performance must be attributed to learning of invariant spatial contexts and not absolute target

location likelihoods. The spatial locations of the targets in the different contexts were counter-balanced across infants, such that Old target locations were New target locations for other infants and vice versa. The distracter locations in each configuration were randomly sampled from eight possible locations, including the target locations used in other configurations. The experiment consisted of 48 trials (4 contexts × 12 trials each) presented as six blocks of eight randomized trials (two per context).

All infants were tested individually in a quiet room, seated at a distance of 60 cm from the monitor on their caregiver's lap. A 5-point calibration sequence (the four corners and center of the screen) was used to obtain the infant's point-of-gaze. The looming calibration stimulus was then presented again in the four corners to validate the accuracy of the calibration. If fewer than four points were accurately calibrated, the sequence was repeated. Average deviation was 1.9° (SD = 1.3°), suitable for assessing eye movements within the specified areas of interest.

Following successful calibration, a colorful attention-grabbing stimulus was presented to draw infants' fixation to the center of the screen. After ensuring fixation, the experimenter manually initiated the first trial to present a search array of four shapes (see Figure 1). After 2 seconds, the shapes appeared to flip over, revealing the face target and three distracters. The infant had 2 seconds to search, after which the target became dynamic for 2 seconds to attract the infant's attention. After a brief pause of 500 ms, a between-trial attention-grabbing stimulus was presented to center infants' fixation before the next trial began. A break was inserted every two blocks, in the form of a 10-second video clip (Elmo from *Sesame Street*). The entire experiment lasted approximately 8 minutes.

2.4 | Data analysis

Eye movements were separated into discrete fixations using a temporal filter of 80 ms and a spatial filter of 150 pixels (equal to 3.63° visual angle). Areas of interest were uniformly delineated around the eight possible target and distracter locations. Fixations that landed in the target AOIs were coded anticipations if they were initiated in the first 2 seconds of the trial (before the target appeared) or reactions if initiated in the remaining 4 seconds of the trial (while the target was visible). The following dependent variables were computed to measure differences in infants' responses to Old and New contexts: (1) Mean saccadic RT latency to orient to the target, and slope of the

RT latency function, calculated separately for each infant across 12 trials per condition; (2) mean number of visits to distracters prior to orienting to the target, and slope of the visits to distracters function, calculated separately for each infant across 12 trials per condition; (3) mean number and proportion of trials in which the target was anticipated; and (4) mean proportion of looking time at the target vs. the distracters prior to target onset. Previous experiments have taken decreased latencies and higher rates of anticipatory looking as evidence of gains in spatio-temporal knowledge (e.g., Amso & Johnson, 2006; Kirkham et al., 2007; Markant & Amso, 2013; Tummeltshammer & Kirkham, 2013).

3 | RESULTS

3.1 | Visual search for targets

Mean RT latencies to orient to the targets were compared in an Age (6 months, 10 months) by Context (Old, New) mixed ANOVA. Results show a significant main effect of Context, $F(1, 44) = 14.02$, $p = .001$, $\eta_p^2 = 0.242$, and no effect of Age ($F(1, 44) = 0.39$, $p = .534$, $\eta_p^2 = 0.009$) or Age by Context interaction ($F(1, 44) = 1.28$, $p = .264$, $\eta_p^2 = 0.028$). Figure 2 shows that infants were faster to orient to the targets in Old contexts than in New contexts, $t(45) = 3.85$, $p < .001$.

To evaluate changes in RT latency due to learning the spatial co-variation of each display, we generated RT latency functions separately for each infant for each condition and compared their slopes. One-sample t tests comparing slopes to 0 showed that mean latencies decreased across trials in Old contexts ($t(45) = 1.93$, $p = .060$) and increased across trials in New contexts ($t(45) = 2.19$, $p = .034$). An Age (6 months, 10 months) by Context (Old, New) ANOVA resulted in a significant main effect of Context, $F(1, 44) = 8.76$, $p = .005$, $\eta_p^2 = 0.166$, and no effect of Age ($F(1, 44) = 0.89$, $p = .351$, $\eta_p^2 = 0.009$) or interaction ($F(1, 44) = 0.08$, $p = .785$, $\eta_p^2 = .002$). Infants' RT latencies to the targets decreased more substantially across blocks (i.e., more negative slopes) when presented in Old contexts than in New contexts, $t(45) = 3.03$, $p = .004$.[1] Further, the intercepts of infants' RT latency functions were compared in the same Age (6 months, 10 months) by Context (Old, New) ANOVA, which yielded no significant effects or interactions (all $p > .352$). Taken together, these results confirm that the differences in infants' saccadic RT latencies to the targets within Old and New contexts emerged through learning of the relevant spatial
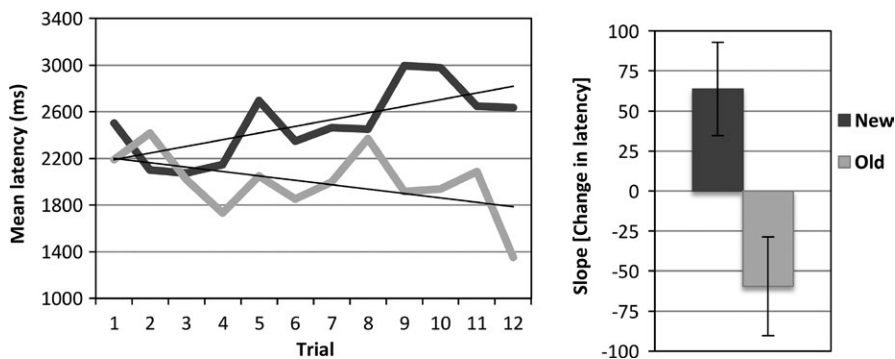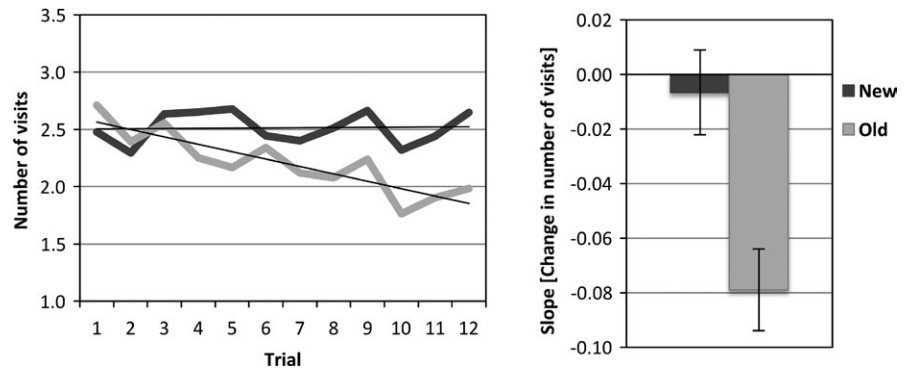

FIGURE 2 Mean RT saccadic latency to targets in Old and New contexts across trials (left), as well as mean change in RT latency (i.e., slope of latency functions) (right)

FIGURE 3 Mean number of visits to distracters prior to visiting the target in Old and New contexts across trials (left), as well as mean change in number of visits (i.e., slope of visit functions) (right)

configurations and resulted from applying that knowledge to more efficiently guide target search.

3.2 | Search efficiency

The efficiency of infants' visual search was quantified as the number of visits they made to distracter locations before orienting to the target (e.g., fixations to distracters A, B, B, C, and A would be recorded as four separate visits). Mean numbers of visits were compared in an Age (6 months, 10 months) by Context (Old, New) mixed ANOVA, which showed a significant main effect of Context, $F(1, 44) = 6.26$, $p = .016$, $\eta_p^2 = 0.124$, and no effect of Age ($F(1, 44) = 0.53$, $p = .473$, $\eta_p^2 = 0.012$) or Age by Context interaction ($F(1, 44) = 1.90$, $p = .175$, $\eta_p^2 = 0.041$). Figure 3 shows that infants visited fewer distracter locations when searching for targets in Old contexts than in New contexts, $t(45) = 2.61$, $p = .012$.

To examine changes in search efficiency across trials, we generated functions for the numbers of visits to distracters in each condition separately for each infant and compared their slopes. One-sample t tests comparing slopes to 0 showed that the mean numbers of visits to distracters decreased significantly across blocks in Old contexts ($t(45) = 5.29$, $p < .001$), but not in New contexts ($t(45) = 0.42$, $p = .674$). An Age (6 months, 10 months) by Context (Old, New) mixed ANOVA resulted in a significant main effect of Context, $F(1, 44) = 13.59$, $p = .001$, $\eta_p^2 = 0.236$, and no effect of Age ($F(1, 44) = 0.01$, $p = .918$, $\eta_p^2 < 0.001$), or interaction ($F(1, 44) = 1.83$, $p = .184$, $\eta_p^2 = 0.040$). Infants' numbers of visits to distracters decreased more substantially across blocks (i.e., more negative slopes) in Old contexts compared to New contexts, $t(45) = 3.78$, $p < .001$. The intercepts of infants' visit functions were also compared in an Age by Context mixed ANOVA, which yielded no significant effects or interactions (all $p > .329$). Taken together, these results indicate that infants' speeded search for targets embedded in Old compared to New contexts resulted from increased search efficiency, as infants looked less at distracters and needed fewer visits to locate the targets.

3.3 | Target anticipation

Additional evidence for contextual cueing comes from the target anticipation interval in the first 2 seconds of the trial (i.e., prior to target onset). We analyzed the proportion of trials in which infants anticipated the target prior to its onset (i.e., had a negative latency) as well

as the proportion of total looking time spent at the upcoming target's location. Mean proportions of trials in which infants anticipated the target were compared in an Age (6 months, 10 months) by Context (Old, New) mixed ANOVA. Results showed a significant main effect of Context, $F(1, 44) = 13.24$, $p = .001$, $\eta_p^2 = 0.231$, a marginal effect of Age, $F(1, 44) = 3.63$, $p = .063$, $\eta_p^2 = 0.076$, but no Age by Context interaction ($F(1, 44) = 0.79$, $p = .378$, $\eta_p^2 = 0.018$). Infants anticipated more targets in Old contexts than in New contexts, $t(45) = 3.74$, $p = .001$, as shown in Figure 4.

Anticipatory looking to the target location as a proportion of total looking time during the first 2 seconds of the trial (i.e., prior to target onset) was compared in an Age (6 months, 10 months) by Context (Old, New) mixed ANOVA. Results showed a significant main effect of Context, $F(1, 44) = 7.08$, $p = .011$, $\eta_p^2 = 0.139$, as infants looked longer at the upcoming target location in Old contexts than in New contexts, $t(45) = 2.75$, $p = .008$. A main effect of Age was also present, $F(1, 44) = 5.23$, $p = .027$, $\eta_p^2 = 0.106$, and reflected longer looking at the upcoming target location by 6-month-olds compared to 10-month-olds. There was no interaction among Age and Context ($F(1, 44) = 0.68$, $p = .414$, $\eta_p^2 = 0.015$), indicating that 6-month-olds looked longer than 10-month-olds during the anticipatory time window in both Old and New contexts. These results are displayed in Figure 5.
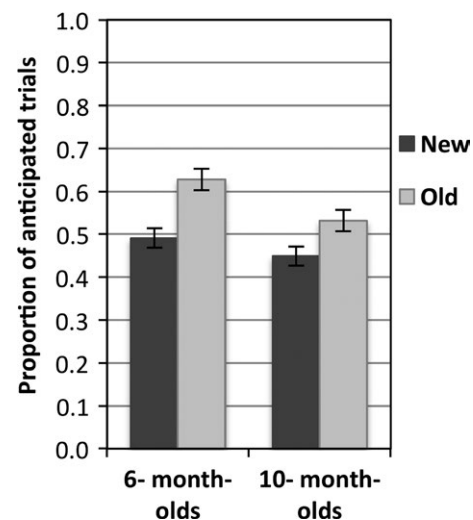

FIGURE 4 Mean proportion of trials in which infants anticipated the target (i.e., had a negative latency)
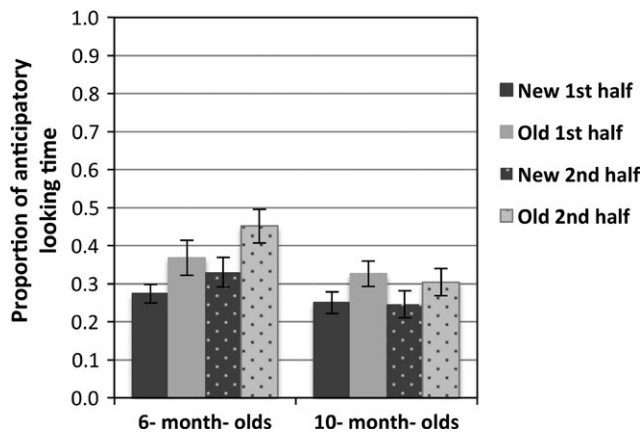
FIGURE 5 Mean anticipatory looking time to the upcoming target location as a proportion of total looking time during the first 2 seconds of the trial (i.e., prior to target onset)

In order to better understand the main effect of Age on anticipatory looking time, we compared performance on the first and second halves of experimental trials, reasoning that the effect may reflect differences in task engagement across the two age groups. An Age (6 months, 10 months) by Context (Old, New) by Trial Number (Trials 1–6, Trials 7–12) mixed ANOVA showed a main effect of Context, $F(1, 44)$ = 8.93, p = .005, η_p^2 = 0.169, a main effect of Age, $F(1, 44)$ = 6.77, p = .013, η_p^2 = 0.133, and an Age by Trial Number interaction, $F(1, 44)$ = 3.89, p = .055, η_p^2 = 0.081. We followed up this interaction with separate post-hoc Age by Context ANOVAs for the first and second halves of trials, which revealed a significant main effect of Age in the second half of trials, $F(1, 44)$ = 10.55, p = .002, η_p^2 = 0.193, and no effect of Age in the first half, $F(1, 44)$ = 0.86, p = .358, η_p^2 = 0.019. Six-month-olds had longer anticipatory looking times to the targets during the second half of the experiment than did 10-month-olds. In summary, we found that the proportion of anticipated trials did not differ by Age, and that the main effect of Age on looking time to the upcoming target location was not present in the first half of the experiment and was similar for both *Old* and *New* contexts. However, we did find that 10-month-olds became less engaged with waiting at the empty target location as the task went on.

4 | DISCUSSION

The visual environment is highly structured, containing redundancies and regularities that may serve to reduce its complexity and constrain visual processes such as object recognition and search. For young infants, this structure may be especially important for establishing priors, resolving ambiguity, and developing stable representations of the visual world. Our results indicate that both 6- and 10-month-old infants are able to orient attention using rapidly acquired top-down knowledge about the structure of the visual environment.

We found clear evidence of infants' sensitivity to visual context, demonstrated by significant differences in their visual behavior when search arrays were repeated or newly generated. Infants oriented

faster to target locations when targets appeared in Old contexts compared to New contexts, and their RT latency functions had more negative slopes, indicating that visual search times became faster across trials within Old rather than New contexts. In their original contextual cueing study, Chun and Jiang (1998) reported that adult participants responded to targets by pressing a button 71 milliseconds faster in Old compared to New contexts during the second half of their experiment; in comparison, we report a contextual cueing advantage of 486 milliseconds on infants' saccadic latencies to targets. Certainly differences in button press and eye movement RTs and a substantial increase in motor control from infancy to adulthood may have contributed; however, it may also be the case that infant viewers, who have slower orienting, processing speeds, and limited working memory capacity, profit more dramatically from the support of stable contextual cues than adults do.

Infants visited fewer distracter locations when searching for targets in Old compared to New contexts, and their search efficiency improved within Old contexts, but not within New contexts. Moreover, infants anticipated the target (i.e., had negative latencies) on a greater number of trials and had higher proportions of looking to the upcoming target's location, rather than the distracter locations, prior to the target's onset. A few studies have indicated that cued or prioritized attention leads to better encoding and recognition memory in infancy (e.g., Amso & Johnson, 2006; Markant & Amso, 2013; Wu & Kirkham, 2010). Thus, contextual cues may play a facilitative role in early learning through the selective deployment of visual attention.

It is relevant to discuss what infants are precisely learning from contextual regularities. Do repeated visual contexts promote learning of adaptive saccade patterns in the service of efficient search, or do contextual cues trigger covert attention processes (or perhaps both)? In Experiment 5 of their adult study, Chun and Jiang (1998) found that contextual cueing benefits persisted even when search arrays were flashed so rapidly that participants did not have time to make multiple eye movements, providing evidence that the cueing effects were not simply due to procedural learning of saccade patterns. In our experiment, search arrays were presented for 2 seconds before the targets were revealed, and our saccadic RT latency data show that even in repeated contexts, infants used nearly the full 2 seconds to orient to the target. Further, our search efficiency data show that infants visited an average of 2.5 distracters before orienting to the target, even after repeated exposure to the Old contexts. Thus, infants executed multiple eye movements during their searches and it is unlikely that they acquired a simple stimulus–response association or saccade program.

We found very few differences between the 6- and 10-month-old groups on any of the measures of visual contextual cueing effects. Six- and 10-month-olds differed marginally in the proportion of trials on which they anticipated upcoming targets, and 6-month-olds did look *longer* at the upcoming target's location (as a proportion of total looking to targets and distracters) than 10-month-olds did. As shown in Figure 4, these age-related differences did not differ according to context; thus they are likely to reflect more general differences in the speed of eye movements and visual attention shifting between 6- and 10-month-olds (e.g., Ross-Sheehy, Schneegans,

& Spencer, 2015). Further, the follow-up analysis comparing first and second trial halves showed that the difference in anticipatory dwell times between age groups was not present in the first half of the experiment, but rather emerged toward the end. This indicates that 10-month-olds were less engaged in anticipating the targets as the task went on, perhaps due to faster or more robust learning than 6-month-olds. Younger infants may also have represented the timing of the target's onset with greater uncertainty, waiting at the expected location for confirmation while older infants continued scanning or disengaged. This is consistent with existing work demonstrating that younger infants anticipate with greater temporal variability (e.g., Canfield et al., 1997; Rose, Feldman, Jankowski, & Caro, 2002). Anticipatory looking as evidence of visual prediction or learning tends to be found in two domains: in repetitive spatiotemporal sequences of visual stimuli where infants demonstrate motor learning of simple saccade patterns (e.g., Canfield et al., 1997; Haith, Hazan, & Goodman, 1988; Reznick, Chawarska, & Betts, 2000) and in action sequences, such as reaching or grasping, where infants show sensitivity to the motion trajectory or end state of the action (e.g., Ambrosini et al., 2013; Falck-Ytter, Gredebäck, & von Hofsten, 2006; Rosander & von Hofsten, 2011). While these studies offer insight into developmental patterns (e.g., anticipatory saccades seem to increase in number and precision with age), they do not require infants to use top-down memory-based representations to generate anticipatory looks.

With respect to all other measures (i.e., latency, search efficiency), 6- and 10-month-olds demonstrated similar performance and equivalent benefits of applying top-down contextual knowledge to guide visual attention and search. However, we do not take this result as evidence that top-down attention orienting is mature and undergoes no significant development in the second half of the first year. Rather, we believe that infants' success on the contextual cueing task signifies an adaptive fit between the demands of the visual search and the component skills involved in top-down control that are refined in the first postnatal year. The framework proposed by Amso and Scerif (2015), which situates visual attention development within the emerging functionality of hierarchically organized visual pathways, describes the development of lower-level visual processing as feeding forward into higher-level regions and acting as a catalyst for top-down attention to exert its influence in managing the increasingly complex visual input. In that sense, top-down attention does not arise at a particular timepoint, but rather is strengthened throughout development by interactions between feedforward and feedback visual pathways. Here, rudimentary learning, attention, and memory skills enabled infants to organize information about color, form, and spatial covariation into higher-order contexts, which could be retrieved to guide their visual responses. Had the task incorporated inputs to which the two groups differed in their sensitivities (or necessitated actions for which they had different capabilities), then we may have seen developmental differences between 6- and 10-month-olds as a function of the system's capacity to manage those demands.

Thus, the current demonstration of contextual cueing in 6- and 10-month-old infants lays the foundation for important and exciting avenues for future research, particularly in determining whether there is an adaptive fit between the contextual demands of their environment and the infant system's ability to manage that complexity. For example, rather than storing visual contextual information in extensive detail, infants may encode salient features or subsets of features to use as cues (see Brady & Chun, 2007; Brockmole, Castelhano, & Henderson, 2006 for work on local vs. global contextual cueing in adults). In their experiment on target and spatial location associations, Bertels and colleagues (2017) found significant correlations among the distances of certain distracters from the target and the size of infants' familiarity preference for repeated arrays, which they took as evidence of more local or feature-based learning. The demands placed on attention and memory systems are likely to impact infants' performance on search tasks involving more complex or conflicting elements (e.g., Gerhardstein & Rovee-Collier, 2002; Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2004). Similarly, while Chun and Jiang (1998) focused on spatial layout as a proxy for visual context, they noted that the content, identity, and features of component objects play an important role in defining the global visual context. Our study preserved the colors and shapes of component items across spatial layouts to provide infants with multiple redundant cues; future studies may consider isolating each of these visual features to assess their contribution to establishing visual context.

The present study has shown that 6- and 10-month-old infants can successfully coordinate learning, memory, and selective attention skills to support the extraction and use of contextual regularities in facilitating visual search. The results offers three noteworthy contributions to our understanding of top-down control and its development: first, that top-down knowledge may influence visual behavior at an earlier stage than expected; second, that the specific time point at which top-down attention emerges is not as meaningful as whether there is an adaptive fit between the task demands and the component skills that support task performance throughout development; and third, that the targets of visual selective attention are both derivatives and determinants of rapid learning, which cyclically shape representations of the structured environment.

ENDNOTE

[1] For the interested reader, we also averaged latencies at four time points (i.e., Trials 1–3, 4–6, 7–9, and 10–12) in order to conduct a 2(Age) × 2(Context) × 4(Time) repeated measures ANOVA. A total of 42 infants provided latencies at all four time points in both conditions to be included. Results show a significant main effect of Context, $F(1, 40) = 8.33$, $p = .006$,

η_p^2 = 0.172, a marginal effect of Time, $F(3, 120) = 2.31$, $p = .079$, η_p^2 = 0.055, and a Context × Time interaction, $F(3, 120) = 2.66$, $p = .052$, η_p^2 = 0.062, all consistent with the slopes analysis reported above. No other effects were present.

REFERENCES

Adler, S.A., & Orprecio, J. (2006). The eyes have it: Visual pop-out in infants and adults. *Developmental Science*, 9, 189–206.

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: Anticipatory gaze and motor ability in infancy. *PLoS ONE*, 8, e67916.

Amso, D., & Johnson, S.P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, 42, 123–1245.

Amso, D., & Johnson, S.P. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy*, 13, 675–686.

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, 16, 606–619.

Bertels, J., San Anton, E., Gebuis, T., & Destrebecqz, A. (2017). Learning the association between a context and a target location in infancy. *Developmental Science*, 20, e12397.

Biederman, I., Mezzanotte, R.J., & Rabinowitz, J.C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143–177.

Brady, T.F., & Chun, M.M. (2007). Spatial constraints on learning in visual search: Modeling contextual cueing. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 798–815.

Brockmole, J.R., Castelhano, M.S., & Henderson, J.M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 699–706.

Brockmole, J.R., & Henderson, J.M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, 13, 99–108.

Butcher, P.R., Kalverboer, A.F., & Geuze, R.H. (2000). Infants' shifts of gaze from a central to a peripheral stimulus: A longitudinal study of development between 6 and 26 weeks. *Infant Behavior and Development*, 23, 3–21.

Canfield, R.L., Smith, E.G., Brezsnyak, M.P., Snow, K.L., Aslin, R.N., Haith, M.M., … Adler, S.A. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, 62, i–160.

Chun, M.M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28–71.

Chun, M.M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, 10, 360–365.

Colombo, J., Ryther, J.S., Frick, J.E., & Gifford, J.J. (1995). Visual pop-out in infants: Evidence for preattentive search in 3- and 4-month-olds. *Psychonomic Bulletin and Review*, 2, 266–268.

Davenport, J.L., & Potter, M.C. (2004). Scene consistency in object and background perception. *Psychological Science*, 15, 559–564.

Dixon, M.L., Zelazo, P.D., & De Rosa, E. (2010). Evidence for intact memory-guided attention in school-aged children. *Developmental Science*, 13, 161–169.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, 9, 878–879.

Fiser, J., & Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, USA*, 99, 15822–15826.

Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, 81, 194–215.

Goujon, A. (2011). Categorical implicit learning in real-world scenes: Evidence from contextual cueing. *Quarterly Journal of Experimental Psychology*, 64, 920–941.

Goujon, A., & Fagot, J. (2013). Learning of spatial statistics in nonhuman primates: Contextual cueing in baboons (Papio papio). *Behavioural Brain Research*, 247, 101–109.

Haith, M.M., Hazan, C., & Goodman, G.S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development*, 59, 467–479.

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.

Johnson, M.H., & Tucker, L.A. (1996). The development and temporal dynamics of spatial orienting in infants. *Journal of Experimental Child Psychology*, 63, 171–188.

Kaldy, Z., & Leslie, A.M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition*, 97, 153–177.

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.

Kirkham, N.Z., Slemmer, J.A., Richardson, D.C., & Johnson, S.P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development*, 78, 1559–1571.

Markant, J., & Amso, D. (2013). Selective memories: Infants' encoding is enhanced in selection via suppression. *Developmental Science*, 16, 926–940.

Markant, J., & Amso, D. (2016). The development of selective attention orienting is an agent of change in learning and memory efficacy. *Infancy*, 21, 154–176.

Merrill, E.C., Conners, F.A., Roskos, B., Klinger, M.R., & Klinger, L.G. (2013). Contextual cueing effects across the lifespan. *Journal of Genetic Psychology*, 174, 387–402.

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science*, 1, 161–184.

Oakes, L.M., Hurley, K.B., Ross-Sheehy, S., & Luck, S.J. (2011). Developmental changes in infants' visual short-term memory for location. *Cognition*, 118, 293–305.

Palmer, T.E. (1975). The effects of contextual scenes on the identification of objects. *Memory and Cognition*, 3, 519–526.

Pelphrey, K.A., Reznick, J.S., Davis Goldman, B., Sasson, N., Morrow, J., Donahoe, A., & Hodgson, K. (2004). Development of visuospatial short-term memory in the second half of the 1st year. *Developmental Psychology*, 40, 836–851.

Reznick, J.S., Chawarska, K., & Betts, S. (2000). The development of visual expectations in the first year. *Child Development*, 71, 1191–1204.

Reznick, J.S., Morrow, J.D., Goldman, B.D., & Snyder, J. (2004). The onset of working memory in infants. *Infancy*, 6, 145–154.

Rosander, K., & von Hofsten, C. (2011). Predictive gaze shifts elicited during observed and performed actions in 10-month-old infants and adults. *Neuropsychologia*, 49, 2911–2917.

Rose, S.A., Feldman, J.F., Jankowski, J.J., & Caro, D.M. (2002). A longitudinal study of visual expectations and reaction time in the first year of life. *Child Development*, 73, 47–61.

Ross-Sheehy, S., Oakes, L.M., & Luck, S.J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science*, 14, 490–501.

Ross-Sheehy, S., Schneegans, S., & Spencer, J.P. (2015). The Infant Orienting With Attention task: Assessing the neural basis of spatial attention in infancy. *Infancy*, 20, 467–506.

Scerif, G., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Developmental Science*, 7, 116–130.

Simons, D.J., & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28, 1059–1074.

Tummeltshammer, K.S., & Kirkham, N.Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16,* 760–771.

Tummeltshammer, K.S., Mareschal, D., & Kirkham, N.Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distracters. *Child Development, 85,* 1981–1994.

Vaidya, C.J., Huger, M., Howard, D.V., & Howard, J.H. (2007). Developmental differences in implicit learning of spatial context. *Neuropsychology, 21,* 497–506.

Wasserman, E.A., Teng, Y., & Brooks, D.I. (2014). Scene-based contextual cueing in pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition, 40,* 401–418.

Wentworth, N., Haith, M.M., & Hood, R. (2002). Spatiotemporal regularity and interevent contingencies as information for infants' visual expectations. *Infancy, 3,* 303–321.

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.

Yoshida, H., Darby, K., & Burling, J. (2011). Cued attention and learning of spatial context in children. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1741–1746). Cognitive Science Society, Austin, TX.