Research Article

What Grasps and Holds 8-Month-Old Infants' Looking Attention? The Effects of Object Size and Depth Cues

Yu Guan and Daniela Corbetta

Department of Psychology, The University of Tennessee, Knoxville, TN 37996-0900, USA

Correspondence should be addressed to Daniela Corbetta, dcorbett@utk.edu

The current eye-tracking study explored the relative impact of object size and depth cues on 8-month-old infants' visual attention processes. A series of slides containing 3 objects of either different or same size were displayed on backgrounds with varying depth cues. The distribution of infants' first looks (a measure of initial attention switch) and infants' looking durations (a measure of sustained attention) at the objects were analyzed. Results revealed that the large objects captured infants' attention first, that is, most of the times infants directed their visual attention first to the largest object in the scene regardless of depth cues. For sustained attention, infants preferred maintaining their attention to the largest object also, but this occurred only when depth cues were present. These findings suggest that infants' initial attention response is driven mainly by object size, while infants' sustained attention is more the product of combined figure and background processing, where object sizes are perceived as a function of depth cues.

1. Introduction

Imagine a mom reading a picture book to her baby. On one page, there are two similar dogs pictured running down an alley. Based on the perspective depth cues provided on the drawing, the leading dog appears bigger and closer to the baby looking at the book, while the other dog, following behind, appears smaller and further away. Which dog will catch the baby's attention first and at which dog will the baby look longer? And what if the depth cues were changed or the sizes of the two dogs were the same? Would the infant produce the same visual attention looking patterns? In this study, we are interested in understanding how infants direct their visual attention to scenes like the one illustrated above, when objects of different sizes are presented on backgrounds with varied pictorial depth cues. In particular, we aim to understand how infants relate object size to depth cues and gain insights on how they integrate figures with their background. By studying how infants visually process information in 2D displays, we can gain better knowledge of the mechanisms underlying infants' cognitive processes.

Visual attention is a complex phenomenon that contains several underlying processes. Cohen was one of the first to examine visual attention in infants. He identified two components of visual information processing that he called "attention getting" and "attention holding" [1]. In Cohen's original 1972 study, looking behavior was assessed while infants were exposed to pairs of checkerboards with checkers varying in size and number. Cohen found that the latency of infants' first head turn toward one checkerboard was mainly determined by the size of the checkers than by their number, while the duration of looking was more determined by the number of checkers than their size. Based on these findings, Cohen suggested a two-process model of infants' visual attention: an *attention-getting* process that determines whether and how quickly infants orient to a stimulus that has been detected, and an *attention-holding* process that determines the duration that infants look at the stimulus once fixated. These two processes can be investigated separately.

Cohen's important contribution to the field of infant visual processing revealed that infants' visual attention is not

a unitary, but a dual process involving different mechanisms of attention [1, 2]. Following on Cohen's footsteps, other researchers continued to investigate different types of attention. Ruff and Rothbart [3] proposed a *visuospatial orienting* response that is specific to infants' initial head and eyes orientation to a newly appearing stimulus. Such initial response is present during most of the first postnatal year of life and is dominated by an orienting/investigative system of attention. The fast maturation in all levels of this system enables infants to deploy rapidly and flexibly their visuospatial orienting abilities to objects in their surrounding from their very early age. Because this process is a fast initial visuospatial orienting response, it requires minimal information processing and is mainly driven by obvious physical features such as salience and size that can be captured easily by a first fixation.

Another component of attention called *sustained attention* is used in contexts when information requires more prolonged scrutiny and further processing [2, 3]. Sustained attention is considered more of a controlled cognitive process dedicated to directing attention to specific object features and details. As a result, this type of attention involves a more protracted voluntary selective engagement of the behavioral system. Furthermore, sustained attention is accompanied by body changes (i.e., in posture, facial expression, and body movements), physiological changes (in heart rate), and brain changes (in EEG waveforms) [4, 5]. For instance, the heart rate drops and stays low when the infant is actively directing attention to the stimulus. Also, evidence has been provided that in sustained attention infants show greater memory encoding [6]. Finally, Colombo et al. [7] suggested that looking duration is negatively correlated with cognitive performance. For instance, short looking infants are more likely to show recognition memory than long looking infants. Also, very long looking duration is not necessary associated to information processing, but rather indicates a low ability to disengage attention [8]. In sum, these studies have demonstrated that active information processing happens during sustained attention, and that this type of attention is accompanied by physiological changes such as a drop in heart rate. Sustained attention is the product of a slower process that requires longer time and more specific information processing in order to explore and combine the details of an object or a scene. This process can be captured by the looking duration and can be influenced by the richness or complexity of the elements in the scene.

In this study, we focus on two related attention processes: the initial switch in attention to a new object on a slide after the infants have already begun orienting to the novel stimulus and the duration and location of sustained attention to a particular region of the scene. We used these two attention processes to better understand how infants visually scan 2D images containing objects of different sizes displayed on backgrounds with varying pictorial depth cues. We aimed to understand whether infants can relate object size to depth cues and whether they can integrate figures with their background. Prior research has shown that object size plays an important role in capturing infant's initial attention. Cohen's study described above showing that infants first oriented to the board with wide checkers is one example of

that. A similar object size effect was reported by Newman and colleagues with 3D objects [9]. These researchers presented two cylinders of different diameters to infants to test their first fixation. They found that infants preferred to shift their first fixation toward the larger cylinder. These studies, however, focused only on the effects of the size features of objects on visual attention. They did not investigate how visual attention to object size was processed in relation to background information. Thus, how infants process visual information in more complex scenes containing multiple objects of various sizes displayed on different backgrounds is fully open to speculation. This is an important question to address as our perception of objects in our surrounding takes place in a world where objects and background provide constant information essential for body orientation and action guidance. Depth cues, in particular, represent important background information in a scene, as they directly relate to objects' size perception. A number of classic studies on size perception in adults have demonstrated that size estimation is more accurate when depth cues are present. For example, Holway's and Boring's [10] classic study showed that when depth cues were progressively removed, a regression occurred in the perceived size of objects that was approaching the visual angle of the object on the retina, not the actual physical size of the object. Other studies revealed the importance of texture gradient as another important depth cue in judging object size [11–13]. Further, a study has indicated that, the more depth cues are available, the fewer adults overestimate the size of objects [14].

We know that infants also rely on pictorial depth cues to perceive object size in 2D displays, but reports show that they do so only from about 5 or 7.5 months old. Since the 1970s, Yonas and his colleagues have systematically studied infants' depth perception using stimuli displaying different pictorial depth cues [15–18]. In one study, infants were presented with two identical 3D toy dolls of the same size suspended in front of a photo of a textured background that depicted a surface receding in depth [18]. Under monocular viewing conditions, 7-month olds reached more for the apparently nearer object, while the 5-month olds did not show any reaching preference to one of the two objects. The researchers concluded that the depth cues background affected the 7-month-old infants' perception of object size but not the 5-month olds. Similar findings were observed in a longitudinal study, confirming that the emergence of sensitivity to pictorial depth cues based on texture gradients and linear perspectives occurred between 5 and 7 months of age [19].

Our interest, in this study, is to examine infants' visual attention processes in relation to stimuli containing figure and background information as in the case of object size and depth cues. Investigating such question will provide important and novel information on how infants process and integrate figures in relation to their background. To explore this research question, we examined how infants directed their visual attention to a series of slides that displayed objects varying in sizes on backgrounds containing different linear perspective depth cues. The number of objects on the slides was always three, so that object number, which was shown by Cohen [1] to affect looking duration, was

controlled. Also, we chose to begin exploring this research question with 8-month-old infants because the research reviewed above clearly indicated by that age, infants are sensitive to depth cues [15–19]. Thus, this age provided us with a good starting point to measure unambiguously how object size is perceived in relation to background depth cues. Furthermore, to specifically assess the relative impact of object size and depth cues on infants' visual-attention processes, we manipulated both the object sizes and the depth cues. Some slides had three objects of different sizes and others had three objects of the same size. We also combined object size with perspective depth cues in such a way that some slides contained depth cues while others did not. Finally, we used eye-tracking to determine on which object infants directed their visual attention first (a measure of initial fixation switch following the stimulus appearance) and on which object they maintained their visual attention the longest (a measure of sustained attention). The use of an automated corneal-reflection system provided a number of advantages: first, it provided accurate information on eye gaze at a relatively high sample rate (50 Hz) and second, it allowed us to determine objectively and accurately to which object the first shift in fixation was directed and on which object looking duration was maintained for the longest. Based on the research reviewed above, we expected that the largest object size would grasp 8-month-old infants' attention first, whether depth cues would be present or absent. Thus, in the conditions when objects are varying in size, the first shift in gaze following the appearance of the slide, should go to the largest object regardless of depth cues because the largest object is the most salient stimulus on the slide. In contrast, when objects are of the same size, we expected infants' first shift in fixation to not be linked to a particular object or location. Our expectations were also that depth cues should be playing more of a role in the sustained attention process as depth cues should be interacting with object size in driving visual attention preferences. We hypothesized that when infants have more time to explore the objects in relation to their background, they should combine object size and background information such as to "hold" their looking attention on the largest object for a longer time particularly when depth cues are present. If this assumption is correct, that the presence of depth cues matters and drives looking preferences to the largest object, then, when objects are of the same size, depth cues should continue to drive infants' looking attention to the object that would appear the largest on the depth background, consistently with Yonas et al. [18] findings on the two dolls experiments. We do not expect observing these looking preferences when depth cues are absent.

2. Method

2.1. Participants. Forty infants (20 females, 20 males) aged 8 months (±7 days) participated in the study. All infants were recruited via formal recruitment mailings and follow-up phone calls. The names and addresses of the participants were obtained through a government-supplied database of births. Additional 45 infants were brought to the laboratory but were excluded from the analyses due to fussiness

FIGURE 1: Experimental setup showing the infant seat, the eye-tracker, the projection screen, and the surrounding enclosure.

($N = 22$), improper eye movement calibration ($N = 7$), or lack of useable eye-tracking data ($N = 16$). According to previous infant studies using remote eye trackers, the attrition rate can range from 33% to 59% [20–22]. Thus the 53% attrition rate in this study is consistent with other studies that relied on this technology. Among our sample that yielded useable data, 35 participants were White and 5 were African Americans. All of them were born full term, and were free of visual impairments. All parents gave consent to have their infant participate in this study and they were given a photograph of their child and a certificate of participation.

2.2. Materials and Stimuli. Figure 1 depicts the experimental set-up. A custom-made infant seat, reclined 10 degrees from vertical, was used to support the participants in front of the testing apparatus. A wide foam strap around their torso provided full trunk support while permitting a full range of motion of the arms and legs.

A small wooden table ($64 \times 38 \times 38$ cm, width $\times$ depth $\times$ height), supporting a remote eye-tracking device (Tobii x50, Tobii Technology, Inc., Danderyd, Sweden) was placed in front of the infant seat such that the eye tracker was located at 60 cm distance from the infant eyes. The eye-tracker used corneal reflection to record where on the slides infants directed their gaze at a frame rate of 50 Hz and with an accuracy of 0.5 degrees. A Dell 3400MP projector (Dell Inc., TX, USA), connected to a computer located in an adjacent room, was positioned under the table out of the view of the child. The projector was used to project stimuli on a large white cardboard screen (102 cm height $\times$ 151 cm width) situated at a distance of 267 cm from the participant. A speaker, connected to the same computer, was behind the projection screen.

Custom designed panels measuring 205 cm (height) $\times$ 155 cm (width) $\times$ 320 cm (length) enclosed the projection area and surrounded both the infant and the projection screen such that extraneous visual distractions were removed. The infant was seated at one end of the panel
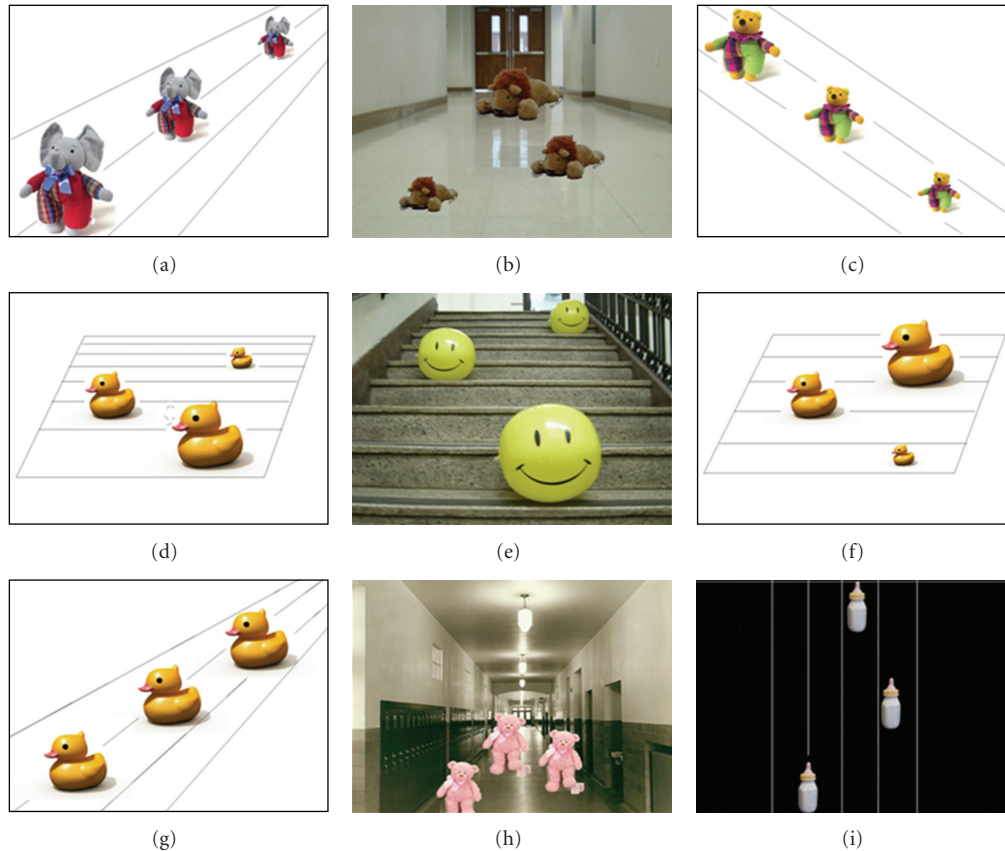
FIGURE 2: Example of slides used for the different conditions. (a)–(f): The 3 objects are of different sizes. (g)–(i): The 3 objects are the same sizes. Columns 1 and 2 display slide exemplars with linear pictorial depth cues and change in texture gradient for the same and different size objects conditions. Columns 3 displays slide exemplars for the no depth conditions with object of same size or different size.

enclosure and the projection screen was located at the other end.

2.2.1. Stimulus Slides. Thirty slides were used as the stimuli. Once projected on the cardboard screen, the slides were 87 cm (height) × 115 cm (width), visual angle: 20.6 degrees vertically, 27.0 degrees horizontally. The stimuli comprised five conditions which corresponded to different combinations of object size variations and depth cues.

Condition 1 (congruent depth cues) included slides that depicted scenes as infants would typically see them in the natural environment, that is, with the object sizes and depth cues being congruent with one another. Three different-size objects (large, medium, small) were scaled gradually and consistently with the background perspective depth cues. Thus, the apparent nearest object was the largest and the apparent more distant object was the smallest (see examples in Figures 2(a), 2(d), and 2(e)).

Slides in conditions 2 and 3 displayed three objects of different sizes scaled either on a background of reversed linear perspective depth cues (condition 2: reversed depth cues), such that the apparent nearest object was the smallest and the apparent furthest object was the largest (see example in Figure 2(b)), or on a background containing no depth cues (condition 3: no depth cues) so that no object appeared

closer or more distant than the others based on background cues (see examples on Figures 2(c) and 2(f)).

Conditions 4 and 5 included slides that displayed three objects of the same size either arranged on a background with perspective depth cues (see examples on Figures 2(g) and 2(h)) or on a background with no depth cues (see example on Figure 2(i)).

All conditions contained a mix of real photos backgrounds and backgrounds with simple lines. The object targets were toys or other infants friendly items such as baby bottles, and the layout of the objects on the slides were also arranged in such a way that the position of certain objects (i.e., the largest object) appeared in different locations of the slides to control for directional looking biases (see Figure 2 for examples of slides used). All stimuli and overlay of the objects on the backgrounds were created in Microsoft Office PowerPoint 2003. All 30 slides were presented in a random order and for two successive rounds with a music background that was synchronized to the slide projection to help capture and maintain infants' visual attention to the slides. Each slide was projected on the screen for a 5-second duration and was immediately followed by a 2-second interslide containing a smiley face at the center of the slide displayed on a uniform black background. This intermediate slide was used to draw the infant's attention back to the center

of the screen before the appearance of the next slide and to allow the researchers to monitor the continued accuracy of the eye signal throughout the entire slide show presentation.

2.3. Procedure. The participants came to the laboratory accompanied by one or two of their parents. Parents signed the consent form after they were explained the goal of the study, procedure, and their questions were answered. Then, the participants were seated in the infant seat facing the projection screen and the eye tracker. One parent was seated close to the infant, but out of his/her view.

Before the experiment started, the participant was presented with an infant video on the projection screen to capture her/his attention. When the infant was looking at the screen and the eye tracker provided a stable signal for both eyes, the video was replaced by stimuli to initiate the calibration procedure. We used a five point calibration procedure. To attract the infant's attention to each calibration point, a computer-generated colored figurine that expanded and contracted over a white background appeared on the projection screen at 5 different locations. These calibration figurines were provided by the Tobii eye-tracker software (Clearview v2.7.1). They appeared for 3 seconds at the top left, top right, bottom left, bottom right, and center of the projection screen. A calibration was considered successful if measures from at least three calibration points were obtained. Otherwise, the procedure would be repeated. The light level during the calibration and the experiment were kept constant to minimize recording errors due to differences in pupil size. The calibration procedure never lasted beyond 5 minutes. If we were unable to calibrate at least three points, the data for that participant were not used.

Once calibration was achieved, all 30 slides with interslides were presented to every participant on the projection screen in a random order. The entire 30 slides presentation took 3 minutes and 28 seconds to play. A second round of the same slides was presented right after the first round, thus, overall, infants saw each slide twice. No recalibration was performed between rounds. Maintenance of calibration accuracy was monitored throughout the entire experiment by checking that the infants' gaze always fell on the smiley faces during the 2-second interslides.

2.4. Data Analyses. All looking data were exported by Clearview v2.7.1 (the software provided by Tobii to run the eye-tracker and analyze the data) in three formats: videos, fixation tables, and gaze plots. We used the gaze plots and videos to identify the objects to which infants switched their visual attention first. For the looking duration analysis, we used the fixation tables.

2.4.1. Initial Attention Switch and Looking Latency to the First Object Fixated. The gaze plots on each slides revealed the order in which infants scanned and successively fixated the scenes on the slides. Because we used eye-tracking and presented a single slide stimulus (as opposed to two in prior studies) right after each interslide with a smiley face at its center, we determined the first fixation switch to the first object attended as follow. Once the stimulus slide appeared, we ignored the very first fixation point on the slide. This is because in 84.13% (sd = 1.24) of the time, the very first fixation point appeared at the center of the slide where the smiley face of the previous inter-slide was located (this very first fixation point appeared at the top of the screen only on 3.79% (sd = 0.59) of the time and at the bottom of the screen only on 2.11% (sd = 1.00) of the times). Although infants already begun orienting to the display when the stimulus slide appeared, we did not consider the very first fixation point on the stimulus slide as the product of an alert and attentive attention switching from the participant. We instead considered the very next fixation point as the infant initial attention switch to a novel object on the stimulus slide. From the gaze plots and videos, we identified the object to which that very next fixation switch was directed and we determined the time elapsed between the first fixation point and that very next object-directed attention switch (latency).

2.4.2. Sustained Attention: Looking Duration to Each of the Three Objects. The fixation table contained the looking time of each fixation as per area of interests (AOIs) defined for each slide. Each slides had 3 AOIs. Each were determined by one object surface and contour, plus an added 1 cm margin surrounding the edges of that object. The looking duration was the accumulated fixations durations within each of these AOIs. Thus, this measure provided the amount of time infants spent looking at each object individually during each 5 seconds slide presentation time.

2.4.3. Criterion for Selection of Useable Data. We only analyzed data from infants who yielded at least 3 seconds of useable eye-tracking data per slides across the two rounds. The looking durations at each object were then normalized by the total looking time at the slide.

3. Results

Our first questions were to determine (1) to which object infants initially switched their visual attention and (2) the time it took infants to direct their attention to such object (latency).

3.1. First Object Visually Attended. Figures 3(a) and 3(b) display the average proportion of times each object was first visually attended after visual attention was switched from its initial center position on the slide. Analyses were performed as a function of object group type (same size versus different sizes) and depth condition. Because these data were frequencies and did not follow a normal distribution, we performed independent non parametric repeated measures ANOVAs on ranks (Friedman tests) on the different groups and conditions to assess specifically whether one object on the slides was first attended visually more frequently than the two others.

3.1.1. Objects of Different Sizes. Figure 3(a) shows that the patterns of initial attention switch to one of the three objects
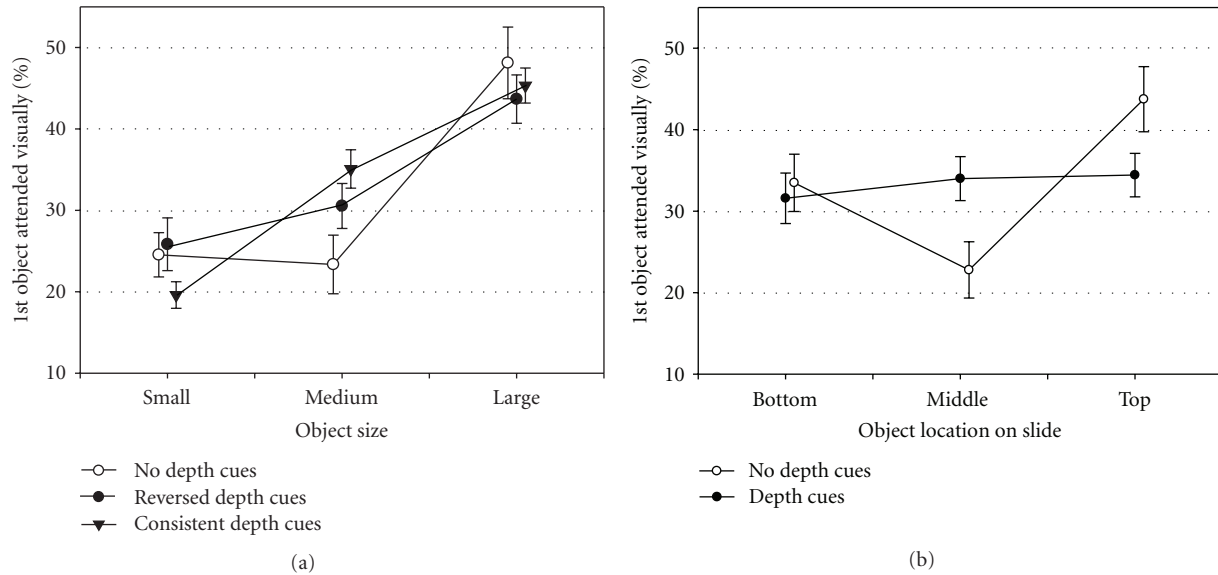
FIGURE 3: Mean percentages (and error bars) of object first visually attended plotted as a function of object type ((a) different size objects; (b) same size objects) and depth cue condition.

on the slides were scaled as a function of the sizes of the object. All infants tended to direct their gaze significantly first and most frequently to the largest object in all three different-size conditions (congruent depth cues: Friedman's $\chi^2(2) = 31.36$, P (exact) = .0001; reversed depth cues: Friedman's $\chi^2(2) = 10.38$, P (exact) = .006; no depth cues: Friedman's $\chi^2(2) = 11.14$, P (exact) = .004).

3.1.2. Objects of Same Sizes. When the three objects were of the same size (Figure 3(b)), only one statistical test led to significant differences in initial direction of visual attention between objects. This was for the no depth cues condition (Friedman's $\chi^2(2) = 11.31$, P (exact) = .004). Figure 3(b) shows that in that condition, when the slide appeared, infants initially switched their gaze more frequently to the objects located at either the top or bottom of the slides, than the object located in the middle of the slide. This result is not surprising. Since the first fixation was in most cases located at the middle of the slide and we removed it, in the absence of size differences between objects, infants were more likely to move their gaze next to one of the two other objects locations on the slide. For the depth cues condition, there was no statistical significant effect.

3.2. Latency to the First Object Visually Attended. This measure captured the time (in msec) infants took to direct their gaze from their first fixation at the center of the slide, when the slide appeared, to the first object attended. These data were averaged as a function of object size and type of depth cues. Because these data met normal distribution and equal variance, they were analyzed using repeated measures General Linear Models. When sphericity assumptions were not met, we used the P values provided by the Greenhouse-Geisser or Huynh-Feldt corrections, depending on whether the Epsilon value was greater or lower than 0.75.

3.2.1. Objects of Different Sizes. When objects on the slides were of different sizes, we ran a 3 depth cues (congruent, reversed, no depth cues) × 3 object sizes (small, medium, large) GLM with repeated measures on the depth and object size factors. Results revealed no main effects of depth cues ($F(2, 76) = 2.747$, $P = .07$), or object size ($F(2, 76) = 1.764$, $P = .178$) on visual latency. Overall, the average looking latency for all objects and depth cues conditions was 632 ms (sd = 34.63).

3.2.2. Objects of Same Sizes. When objects on the slides were of identical sizes, we ran a 2 depth cues (depth cues, no depth cues) × 3 object locations (bottom, middle, top) GLM with repeated measures on the depth and object location factors. Results only revealed a main effect of object location ($F(2, 76) = 5.59$, $P < .005$). There was no main effect of depth cues ($F(1, 38) = 0.001$, $P = .974$). Repeated measure contrasts on object location revealed that infants had greater latencies when directing their gaze to the bottom object first (764 ms) compared to latencies to the middle or top objects first (701 ms and 707 ms, resp., $P < .013$).

3.3. Looking Duration to the Different Objects. Our next question was to determine whether, after the first look at one object, accumulated visual attention was maintained toward that object or directed toward another object (measure of sustained attention). We also wanted to determine whether sustained attention to one object was modulated by background depth cues (a measure of figure/ground integration). Looking durations to each of the 3 objects on the slides were averaged by depth cues condition and were entered into 2 separate repeated measures GLMs depending on the type of object presented (same sizes versus different sizes) after verifying that the data met the requirements of normal
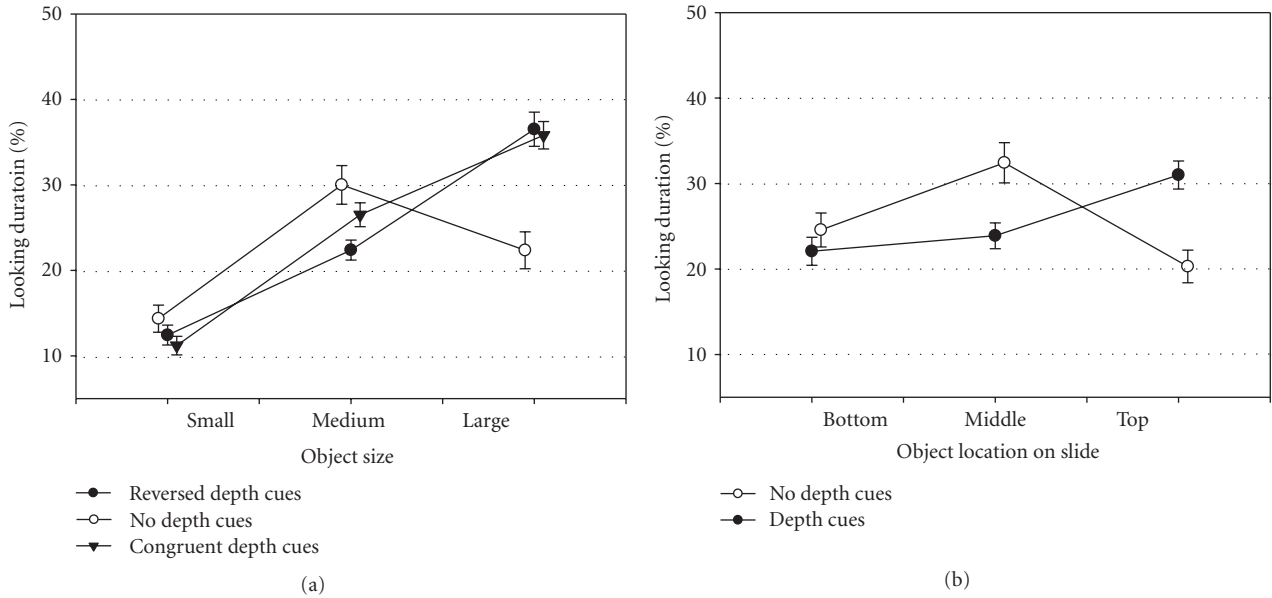
FIGURE 4: Mean percentages (and error bars) of looking duration on each object plotted as a function of object type ((a) different size objects; (b) same size objects), and depth cue conditions.

distribution and equal variance. Again, when sphericity assumptions were not met, we used the P values provided by the Greenhouse-Geisser or Huynh-Feldt corrections, depending to whether the Epsilon value was greater or lower than 0.75.

3.3.1. Objects of Different Sizes. A 3 depth conditions (congruent, reversed, no depth cues) × 3 object sizes (small, medium, large) GLM with repeated measures on the two last factors revealed a main effect of depth cues ($F(2, 76) = 8.066$, $P < .001$), a main effect of object size ($F(2, 76) = 65.082$, $P < .0001$), and a significant depth × object size interaction ($F(4, 152) = 15.236$, $P < .0001$). Figure 4(a), displaying these results, shows that on average, infants looking duration increased with the size of the object attended. Pairwise comparisons revealed that infants looked significantly longer at the medium object compared to the smaller objects ($P < .0001$) and that they looked longer at the largest object compared to the medium one ($P < .002$), but this object effect was moderated by the presence or absence of depth cues ($P < .0001$). Interestingly, sustained attention to the largest object occurred only when there were depth cues in the background (congruent and reversed depth cues). When background depth cues were absent (no depth condition), infants no longer displayed sustained attention to the largest object. This result is particularly notable, because it indicates that the observed looking durations were not purely a proportional response to the size of the object attended, but rather were the combined product of attending the objects in relation to their background.

3.3.2. Objects of Same Sizes. Figure 4(b) displays the average looking duration at each object as a function of depth cues. A 2 depth conditions (depth cues, no depth cues) × 3

object locations (bottom, middle, top) GLM with repeated measures on the two last factors revealed no main effect of depth cues ($F(1, 38) = .015$, $P < .905$) and no main effect of object size ($F(2, 76) = 2.757$, $P < .07$), but a significant depth × object location interaction ($F(2, 76) = 20.180$, $P < .0001$). As for the initial attention switch, this data show that once variations in object size were removed, infants did not display a strong looking preference for any of the objects in particular, but the significant interaction with depth cues adds an interesting twist to these results. Follow-up contrasts revealed that the significant interaction rested right in between the middle and top object. When depth cues were present, infants looked significantly longer at the top object compared to the middle object ($P < .0001$), whereas when depth cues were absent they did not. Interestingly, in the depth cues condition, the top object was the one close to the receding lines or the one located where increased texture gradient depth cues were present. This finding may suggest that infants may have spent more time looking at that object in that condition, because it may have appeared larger to them given the depth cues present. The fact that infants did not show the same looking trend in the no-depth condition suggests again that infants did not just process object and their location in their looking patterns, but integrated figure/ground information.

4. Discussion

According to Cohen's [1] and others [2–8], infants' visual attention of 2D displays contains two distinct processes. First, a visuospatial orienting response which is a fast initial orienting response that requires little information processing and is measured by the direction and latency of the first fixation. Second, sustained attention which is a slower

process that requires some information processing time and is measured by looking duration. Although past research on infants' visual attention has emphasized the roles of object's properties (i.e., size) and scene complexity (i.e., object number) on these attention processes, they did not measure the individual and interrelated effects of object size and depth cues on these two processes. This is the first study assessing the roles of object size and depth cues in 8-month-old infants' orienting response and sustained attention using eye-tracking technology.

Because we used eye-tracking and therefore a method that differed from prior studies that assessed orientation and looking preference between two adjacent slides, our measure of attention orientation was more a measure of attention switching from an initial central position on the slide to a new one where an object was located. We considered that switch as our measure of first fixation. Results showed that, when the three objects were of *different sizes*, infants tended to shift their first fixation toward the largest object no matter where it was located on the slide or whether depth cues were present or not. In contrast, when the three objects were of the *same size*, infants did not reveal a significant preference in attention direction in relation to their first fixation. For the results on looking duration, when the three objects were of *different sizes*, infants spent longer time looking at the largest object on the slide, but they did this only when depth cues were provided in the background. When depth cues were absent, infants' looking preferences for the largest object or any object disappeared. Furthermore and interestingly, in the *same size* conditions, when the depth cues were provided, infants' revealed a looking preference for the object located at the converging/texture gradient receding side of the depth cues. However, again, when depth cues were absent, infants did not display such looking preference for any of the three objects.

Our results on the first fixation switch are in agreement with findings from several prior developmental studies that showed that infants usually tend to direct their visual attention first to the larger objects, because larger objects are more effective attention grabbers [1, 9]. Thus, consistent with Ruff's and Rothbart's [3] suggestion, infants in the first postnatal year will respond to object salience quickly. Such rapid visuospatial orienting response requires little information processing.

The novel aspect of this research resided mostly in extending our understanding of the sustained attention in infants' visual exploration. According to early reports of attention development during the first year, the maturation of the posterior and anterior attention systems increase infants' visual ability such that they can begin to respond to other objects properties, and not just object salience [3]. Furthermore, behavioral research has suggested that attention holding is particularly the product of information processing, and, therefore, can be directly influenced by the richness, details, or complexity of the patterns [23, 24]. For instance, Cohen's finding revealed that the number of checkers determined looking duration [1]. The more checkers were provided in the scene, the longer time infants spent looking at them because infants required longer time to process the

rich information. In our study, we controlled the richness of the foreground information by providing an equal numbers of objects in all slides but we varied the depth cues background information. Interestingly, using these stimuli, our findings demonstrated that the depth information in the background can also influence infants' looking duration preferences in a dynamic way. For instance, when the slides contained objects of *different sizes*, infants' looking patterns preferences were altered depending on whether there were depth cues or not in the background. Specifically, when depth cues were present, they spent longer time looking at the largest object, but this effect dissipated when the depth cues were removed. We think this result is linked to the time infants had to explore and process the information on the slides visually. When depth cues were present in the background, they came into play in the perception process and interacted with object size information. In other words, the longer looking duration enabled infants to combine the size and depth cues information in the scene so as to hold their visual attention on one object only when depth cues were present. Yet, even though during the initial orienting response infants looked first at the largest object in the scene, we found that object size alone could not continue to hold infants' looking attention on the largest object during continued exploration of the slide. Only the interaction of both size and depth cues could hold infants' attention to that more salient object. Thus, the novelty of this study helps us further understand the sustained attention process by showing that richness and complexity of foreground patterns are not the only factors that can hold infants' visual attention. As this study shows, foreground patterns with similar richness, but placed on backgrounds containing varying depth cue information can also contribute to "hold" 8-month-old infants visual attention discriminately.

Another interesting issue that emerged from these data on looking duration concerns the possibility that infants may have been responsive to a certain kind of 2D visual illusion that involved object size and depth cues. The Ponzo illusion has been widely studied in adults and findings have shown that linear perspective depth cues in the background have a powerful effect on size perception. In the Ponzo illusion, two rods of the same size are scaled with linear perspective depth cues, the rod on the converging sides (usually on the top part of the slide) looks larger because the scaling mechanism corrects for the apparently increased depth [25]. In this study, the same object size condition with depth cues background created a similar Ponzo illusion visual effect, that is, the object that was located at the converging/receding side of lines/texture gradient could have been perceived to be larger than the other two objects (see Figures 2(g) and 2(h)). When infants were presented with these particular types of slides (objects same size with background depth cues), we found that their fixation time was longer on the object that was located on the converging texture gradient receding side of the depth cues. In contrast, when depth cues were absent (object size was still the same), infants did not show any looking preference for any given object. These findings share some similarities to the way adults perceive the Ponzo illusion. Although adult research has indicated that the

Ponzo illusion has a powerful effect on adults' perception of size, no developmental study so far has investigated this effect with infants. We believe the findings from this study may provide the first evidence that infants as young as 8-month old can respond to such illusion. The fact that they spent more time looking at the object that should appear larger in the illusion could suggest that they were able to detect the effects created by the depth cues on the object size.

An alternate possible explanation for this particular finding in the same object size condition could be that infants examined more the object near the greater gradient receding lines when depth cues were present because that area contained more information and therefore was more complex to explore. To investigate that possibility, we examined whether infants spent some time looking at the background lines surrounding the different objects on the slides. First, we found that infants hardly explored the lines on the slides background; they spent most of their looking time on the objects *per se*, not the background. Second, their little exploration of the lines was no greater at their receding point than at any other locations on the slides. This suggests that background depth cues were not processed by the infants through the direct perception of the specific depth cues provided on the slides (i.e., lines and other cues) but rather were gathered peripherally as they were directing their attention on the object targets. A recent study has demonstrated indeed that infants are capable of processing peripheral information even when their visual attention is attracted to somewhere else in the visual filed neighborhood [26]. The results of this analysis on the amount of looking directed specifically at the background lines would lead us to infer again, that infants possibly directed greater attention to the top object when objects were of the same size and depth cues were provided because that object may have appeared larger due to the Ponzo illusion, and not necessarily because there was more information to explore in that area of the slide in that condition. But more investigations will be necessary to examine this issue more in depth.

We used 8-month-old infants in this study mainly because previous literature demonstrated that infants become sensitive to pictorial depth cues between 5 to 7.5 months old [15–19]. Thus, the age of 8-month old provided us with a good starting point to measure unambiguously how object size is perceived in relation to background depth cues. And as we expected, infants of this age group were able to combine depth cues information with object size so that their attention was "held" longer by the depth cues on the largest object. In future studies, we are planning to investigate these effects with younger infants to establish the developmental trajectory of these orienting responses and looking durations to figure-ground perception. Particularly, we are interested in assessing whether depth cues will lead to similar looking trends in younger infants and whether younger-aged infants can also show some form of sensitivity to what might be the Ponzo illusion. We speculate that infants younger than 4-month old might not be able to fully integrate figure/ground perception and thus may not respond differentially to the varied pictorial depth cues as the 8 mon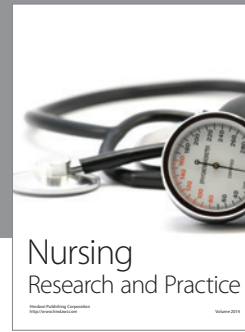ths old in this study did. In addition, infants aged between 5 and 7.5 month might show some variations in their responses to depth cues since this is a transition period in the development of sensitivity to pictorial depth cues. But on average, we anticipate that the influence of the depth cues on object size during sustained attention should become progressively stronger during the first year of life as a result of infants' increased information processing ability.

In conclusion, this research expanded our understanding of infants' initial attention direction and sustained attention when processing a scene visually. We have shown that these attention processes go beyond object salience and complexity and apply to figure/ground perception. We also have shown that 8-month-old infants are well capable of integrating and combining efficiently the foreground and background elements of a scene. Indeed, in our task, infants responded differentially to our 2D picture stimuli depending on whether objects had similar or different sizes but also, and more importantly, whether the background depth cues were present or not.

References

[1] L. B. Cohen, "Attention-getting and attention-holding processes of infant visual preferences," *Child Development*, vol. 43, no. 3, pp. 869–879, 1972.

[2] J. E. Richards, "Infant attention, arousal, and the brain," in *Infant Perception and Cognition: Recent Advances, Emerging Theories, and Future Directions*, L. M. Oakes, C. Cashon, M. Casasola, and D. H. Rakison, Eds., Oxford University Press, New York, NY, USA, 2011.

[3] H. A. Ruff and M. K. Rothbart, *Attention in Early Development: Themes and Variations*, Oxford University Press, New York, NY, USA, 2001.

[4] J. E. Richards and B. J. Casey, "Development of sustained visual attention in the human infant," in *Attention and Information Processing in Infants and Adults*, B. A. Campbell, H. Hayne, and R. Richardson, Eds., Lawrence Erlbaum Associates Press, Hillsdale, NJ, USA, 1992.

[5] G. D. Reynolds and J. E. Richards, "Infant heart rate: a developmental psychophysical perspective," in *Developmental Psychophysiology: Theory, Systems, and Applications*, L. A. Schmidt and S. J. Segalowitz, Eds., Cambridge University Press, Cambridge, UK, 2007.

[6] J. E. Richards, G. D. Reynolds, and M. L. Courage, "The neural bases of infant attention," *Current Directions in Psychological Science*, vol. 19, no. 1, pp. 41–46, 2010.

[7] J. Colombo, D. W. Mitchell, J. T. Coldren, and L. J. Freeseman, "Individual differences in infant visual attention: are short lookers faster processors or feature processors?" *Child Development*, vol. 62, no. 6, pp. 1247–1257, 1991.

[8] J. Colombo, "On the neural mechanisms underlying individual differences in infant fixation duration: two hypotheses," *Developmental Review*, vol. 15, no. 2, pp. 97–135, 1995.

[9] C. Newman, J. Atkinson, and O. Braddick, "The development of reaching and looking preferences in infants to objects of different sizes," *Developmental Psychology*, vol. 37, no. 4, pp. 561–572, 2001.

[10] A. H. Holway and E. G. Boring, "Determinants of apparent visual size with distance variant," *The American Journal of Psychology*, vol. 54, pp. 21–37, 1941.

[11] J. C. Hull, R. T. Gill, and S. N. Roscoe, "Locus of the stimulus to visual accommodation: where in the world, or where in the eye?" *Human Factors*, vol. 24, no. 3, pp. 311–319, 1982.

[12] J. H. Iavecchia, H. P. Iavecchia, and S. N. Roscoe, "The moon illusion revisited," *Aviation Space and Environmental Medicine*, vol. 54, no. 1, pp. 39–46, 1983.

[13] L. Kaufman and I. Rock, "The moon illusion, I," *Science*, vol. 136, no. 3520, pp. 953–961, 1962.

[14] J. W. Meehan and T. J. Triggs, "Magnification effects with imaging displays depend on scene content and viewing condition," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 30, pp. 487–494, 1988.

[15] A. Yonas, W. T. Cleaves, and L. Pettersen, "Development of sensitivity to pictorial depth," *Science*, vol. 200, no. 4337, pp. 77–79, 1978.

[16] A. Yonas, L. Pettersen, and C. E. Granrud, "Infants' sensitivity to familiar size as information for distance," *Child Development*, vol. 53, no. 5, pp. 1285–1290, 1982.

[17] C. E. Granrud, A. Yonas, and L. Pettersen, "A comparison of monocular and binocular depth perception in 5- and 7-month-old infants," *Journal of Experimental Child Psychology*, vol. 38, no. 1, pp. 19–32, 1984.

[18] A. Yonas, C. E. Granrud, M. E. Arterberry, and B. L. Hanson, "Infants' distance perception from linear perspective and texture gradients," *Infant Behavior and Development*, vol. 9, no. 3, pp. 247–256, 1986.

[19] A. Yonas, C. A. Elieff, and M. E. Arterberry, "Emergence of sensitivity to pictorial depth cues: charting development in individual infants," *Infant Behavior and Development*, vol. 25, no. 4, pp. 495–514, 2002.

[20] D. Amso and S. P. Johnson, "Learning by selection: visual search and object perception in young infants," *Developmental Psychology*, vol. 42, no. 6, pp. 1236–1245, 2006.

[21] D. Corbetta, Y. Guan, and J. L. Williams, "Infant eye-tracking in the context of goal-directed actions," *Infancy*, vol. 17, no. 1, pp. 102–125, 2012.

[22] P. C. Quinn, M. M. Doran, J. E. Reiss, and J. E. Hoffman, "Time course of visual attention in infant categorization of cats versus dogs: evidence for a head bias as revealed through eye tracking," *Child Development*, vol. 80, no. 1, pp. 151–161, 2009.

[23] M. L. Courage, G. D. Reynolds, and J. E. Richards, "Infants' attention to patterned stimuli: developmental change from 3 to 12 months of age," *Child Development*, vol. 77, no. 3, pp. 680–695, 2006.

[24] L. B. Cohen, J. S. DeLoache, and M. W. Rissman, "The effect of stimulus complexity on infant visual attention and habituation," *Child Development*, vol. 46, no. 3, pp. 611–617, 1975.

[25] H. W. Leibowitz and H. A. Pick, "Cross-cultural and educational aspects of the Ponzo perspective illusion," *Perception & Psychophysics*, vol. 12, no. 5, pp. 430–432, 1972.

[26] F. Farzin, S. M. Rivera, and D. Whitney, "Spatial resolution of conscious visual perception in infants," *Psychological Science*, vol. 21, no. 10, pp. 1502–1509, 2010.