Young Infants' Eye Movements over "Natural" Scenes and "Experimental" Scenes

Marc H. Bornstein, **Clay Mash**, and **Martha E. Arterberry**
Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services and Colby College

Abstract

Eye movements of 30 4-month-olds were tracked as infants viewed animals and vehicles in "natural" scenes and, for comparison, in homogeneous "experimental" scenes. Infants showed equivalent looking time preferences for natural and experimental scenes overall, but fixated natural scenes and objects in natural scenes more than experimental scenes and objects in experimental scenes and shifted fixations between objects and contexts more in natural than in experimental scenes. The findings show how infants treat objects and contexts in natural scenes and suggest that they treat more commonly used experimental scenes differently.

Keywords

infants; eye movements; natural scenes; object perception; context

When we open our eyes, we see the world in scenes, and the visual system and visual cognition evolved phylogenetically and develop ontogenetically to process natural scenes. The study of adult natural scene processing (especially via eye movements) has garnered attention in recent years (e.g., Bartels & Zeki, 2004; Li, VanRullen, Koch, & Perona, 2002; Rayner, 1998; Rayner, Smith, Malcolm, & Henderson, 2009; Rolls, Aggelopoulos, & Zheng, 2003). By contrast, much less is known about infants' processing of natural scenes. How do infants scan natural scenes and acquire knowledge about objects and their contexts in natural scenes? In the present study, we explore infants' perception of natural scenes. We presented infants with animals in nature scenes (e.g., a monkey on a tree log) and vehicles in residential scenes (e.g., a SUV in a parking lot). To examine infants' perceptions closely we tracked their eye movements as they scanned the scenes. To explore the generality of our findings we tested infants with multiple exemplars nested in the 2 categories (animals and vehicles).

In natural scenes, objects and their contexts are congruent. Thus, infants' perceptions of and cognitions about objects almost always occur in normally congruent contexts. Part of infants' knowledge about objects includes learning about the typical, probable, expected, or congruent contexts in which objects appear. Notably, however, much of the work in infant perception strips objects of their natural scene contexts. Even a cursory survey of the

perception literature reveals that researchers predominantly display objects in homogeneous or neutral (black, gray, or white) contexts. That is, objects in experiments are very often presented devoid of their natural contextual information, and object perception studies use mostly abstracted, idealized, or static stimuli, presumably to segregate and study function. Theory and research vary about the relative advantages and disadvantages of object processing in natural scenes versus isolated contexts (Bartels & Zeki, 2004; Biederman, 1981, 1987; Davenport & Potter, 2004; Palmer, 1975; Righart & de Gelder, 2006; Williams & Weisstein, 1978). We do not conceive of objects embedded in natural scenes and isolated objects as directly comparable. In this study, we focus on infants' perceptions of objects and contexts in natural scenes. For comparison purposes, we studied infants' perceptions of the same objects isolated from their natural scene contexts. For these experimental scenes, we cut the same animals and vehicles out of their natural contexts and imported them into homogeneous contexts (a plain white background). The paucity of research on natural scene perception in infancy complicates attempts to evaluate and interpret research on infant perception and perception of objects. For this reason, we included a condition to represent the experimental circumstance that has been studied more extensively. By including an experimental scene condition, we are better able to evaluate the patterns of results obtained in analysis of data from the focal natural scene condition.

Our main analyses addressed three questions: *What is the pattern of infant allocation of attention to objects and their contexts in natural scenes? Do patterns of infant attention vary for different categories of stimuli, specifically animal natural kinds versus vehicle designed artifacts? Do they vary as a function of the object-context relation?*

Thirty infants (M age = 127.40 days, SD = 9.65, range = 111–150; 13 females) participated: 14 infants in the natural scene condition and 16 infants in the experimental scene condition. An additional 6 infants began the procedure, but they were not included due to fussiness (n = 2) or equipment/measurement failure (n = 4). All infants were term, healthy at birth and at the time of testing, and represented families of middle to upper-middle SES on the Hollingshead (1975) Four-Factor Index of Social Status (Bornstein, Hahn, Suwalsky, & Haynes, 2003).

The total stimulus array consisted of 36 full-color digitized images (Figure 1) divided into 2 sets of stimuli corresponding to 2 conditions: natural scenes and experimental scenes. Color photographs of 9 animals and 9 vehicles were first obtained in their natural environments, such as fields, forests, or lakesides and streets, driveways, or parking lots, and these images constituted the natural scenes as they originally appeared. The experimental scenes were created by digitally manipulating the original photographs: The target objects of the scenes were extracted from their contexts using a computer graphics software package and then imported into a homogenous white background leaving the size and other relevant visual features unchanged.

Images in whole subtended 23° high by 29° wide on average, and objects were 19° high by 26° wide on average. The stimulus images varied modestly in size but did not differ between categories (animals: M = 140.0 cm^2, SD = 39.0; vehicles: M = 148.4 cm^2, SD = 18.1), t(16) = 0.59, *ns*, or (to promote comparability) conditions. Stimulus images were also uniform in mean power spectral density across conditions at both low (0.03–4.95 cy/cm), F(1, 32) = 0.01, *ns*, and high (11.55–16.50 cy/cm), F(1, 32) = 0.75, *ns*, spatial frequency bands.

An Applied Science Laboratories Model 504 infant eye-tracking system was used to record infants' eye movements using infrared corneal reflection relative to coordinates on the stimulus plane continuously at 60 Hz (recording point of gaze every 17 ms). An Ascension Technologies electromagnetic motion tracker, attached to an infant head band, corrected

camera angles for spontaneous head movements that exceeded the frame limits of optical tracking. Eye movement recording was synchronized with stimulus presentation using GazeTracker software that was running on a second microprocessor. Stimuli were presented on a 21- by 29-cm Panasonic video monitor.

Infants were randomly assigned to one of the 2 conditions. Infants were 90 cm in front of the monitor on which the stimuli were displayed, and the transmission component of the head movement tracking system was situated to the rear. The eye camera was located beneath the stimulus monitor in the same depth plane as the stimulus monitor's screen. The eye tracking system was calibrated for each infant individually following conventions in Gredebäck, Johnson, and von Hofsten (2010, p. 3) by presenting rotating red 1.3° plus signs in the upper-left and lower-right corners of an otherwise uniform field.

Each infant saw all 18 stimuli from the assigned condition, 9 animals and 9 vehicles, in one of two random orders with the constraint that no more than two images from the same category (i.e., animals or vehicles) appeared consecutively. Between trials, a uniform field of 16 black 2.5° plus signs maintained infant attention toward the stimulus screen without systematically biasing fixation toward any particular region of the display. Each trial was initiated by a key press when the infant was judged to be looking toward the display. Stimuli were presented for 10 s each.

For each trial, fixations of 200 ms or more were plotted directly on the stimulus image with guide lines drawn 1° in radius around each fixation point using the GazeTracker software package (see Gredebäck, et al., 2010, pp. 9, 12). Plots were coded by visual inspection of coders attending to two areas of interest (AOI). Fixations were classified as "object" if their corresponding guide lines fell on the object without spanning 1o from the object's outer boundary. Fixations were classified as "context" if their corresponding markers fell in any remaining portion of the scene lying beyond 1° from the object boundary. (For additional details about these parameters, see Bornstein, Mash, & Arterberry, in press.) Twenty percent of the sessions were coded by a second observer, and ratings coincided on 98% of the individual trials.

On average, infants contributed 15 trials of usable data (range = 9–18). An initial analysis was conducted in line with more conventional preferential-looking procedures to examine whether infants showed any baseline preference for a particular condition. Total looking time was calculated by summing over all fixations of object and context together for each trial. Mean looking times did not differ between animals ($M = 17.32$ s, $SD = 13.00$) and vehicles ($M = 17.45$ s, $SD = 12.28$), $F(1, 28) = 0.50$, $ns.$, or between natural scenes ($M = 41.80$ s, $SD = 23.02$) and experimental scenes ($M = 28.80$ s, $SD = 24.85$), $F(1, 28) = 2.19$, $ns.$

Because the objects varied somewhat in size, the number of looks at the object on each trial was weighted by the proportion of the image taken by the object in that scene, and the number of looks at the context was weighted by the remaining proportion of that scene. Analyses were conducted on weighted values; looks falling on boundaries were not included. Frequency and duration of fixations correlated highly (object fixations $r = .96$, context fixations $r = .94$), and parallel analyses of frequencies and durations revealed the same patterns of findings; therefore, only numbers of weighted fixations are reported.

Infants' weighted mean numbers of fixations are presented in Figure 2 by stimulus condition and AOI. Weighted mean numbers of fixations were analyzed in a $2 \times 2 \times 2$ ANOVA with condition (natural scenes, experimental scenes) as a between-infants factor and category (animal, vehicle) and location (object, context) as within-infants factors. The analysis utilized planned contrasts to compare, between conditions, infants' overall looks, object

looks, context looks, and looks by category. Infants viewing natural scenes produced more fixations overall per trial ($M = 2.03$, $SD = 1.19$) than infants viewing experimental scenes ($M = 1.18$, $SD = 0.98$), $F(1, 28) = 4.60$, $p = .041$, $\eta^2 = .14$. Infants fixated objects more often when they appeared in natural scenes than when the same objects appeared in experimental scenes, $F(1, 28) = 6.04$, $p = .020$, $\eta^2 = .18$. Infants fixated contexts similarly between natural scenes and experimental scenes, $F(1, 28) = 1.68$, *ns*. Fixations by category did not differ between conditions, $F(1, 28) = 0.08$, *ns*, indicating that fixation counts were similar between categories when viewing natural scenes and experimental scenes alike.

In addition to the number of fixations by AOI, eye movement patterns were examined by coding each instance in which infants shifted fixation from one AOI to the other (i.e., from object to context or the reverse). Consecutive fixations were more likely within AOIs than between AOIs with the mean number of such shifts ≈ 1 per trial. The condition effect for shift frequency between conditions was significant, $F(1, 29) = 7.39$, $p = .011$, $\eta^2 = .21$; as would be expected, infants who viewed natural scenes shifted fixations between AOIs more frequently ($M = 0.92$, $SD = 0.46$) than did infants who viewed experimental scenes ($M = 0.51$, $SD = 0.37$).

This work is among the first to examine natural scene perception in infants; understanding scene perception is an important objective as indicated by the large amount of adult work done over recent years. The central questions explored in this study addressed infant attention allocation to natural versus experimental scenes and to objects and their contexts in natural versus experimental scenes and whether and how infant attention allocation varies as a function of the category of the object (animals versus vehicles) and, to some degree, the object-context relation. Analyses of infants' fixations revealed that, whether inspecting natural or experimental scenes, infants look at both objects and their contexts. Moreover, infants explore animals and their contexts equally to vehicles and their contexts, and they look overall the same amounts at natural and experimental scenes. However, infants give more individual fixations to natural scenes and to objects in natural scenes than to experimental scenes and to objects in experimental scenes and infants compare (shift fixations between) objects and their context more when objects appear in natural scenes than in experimental scenes.

The two conditions in this study are not directly comparable; the natural scenes contain more information than the experimental scenes, of course. Because most research looks at infants' perceptions of objects in experimental scenes, and less at infants' perceptions of natural scenes and objects in natural scenes, we included both. Including experimental scenes facilitates comparisons of our methods with those employed in other research studying comparable object-context relations. The obvious confound of differential amounts of information in stimuli in the two conditions limits the interpretation of differences between the two groups. Because of the predominant nature of testing infant perception using "experimental" scenes, we made the tentative and heuristic natural versus experimental scenes comparison.

For the young infant, learning about objects and their contexts presents a formidable challenge, and opinions on how infants perceive objects in contexts vary. One contention has been that, in navigating a multidimensional world that is constantly changing, infants monitor the environment and differentiate object-context relations by deploying their attention selectively and flexibly. On this argument, it has been asserted that "virtually all learning during infancy is … independent of context" (Nadel, Willner, & Kurz, 1985, p. 398). However, empirical research with infants as young as 3 months of age has challenged this position, and it has been counter-claimed that, when infants monitor objects in their environment, context cues are salient. For example, Colombo, Laurie, Martelli, and Hartig

(1984) observed that surrounding segments influence infants' pattern discrimination (the configural superiority effect), and Haaf, Lundy, and Coldren (1996, p. 96) observed that infants habituate to a focal cue more quickly when its context is constant than when its context varies, so it appears that "infants … attend to and encode background context information while encoding central stimulus cues." Thus, context manifestly affects infant attention and perception. Rovee-Collier's work shows that a change of context can result in failure to display previously learned contingent behaviors (Butler & Rovee-Collier, 1989; Rovee-Collier & DuFault, 1991). In this study, we show that context also matters as infants eye natural scenes.

Extending from our findings, it is likely that, as human infants encounter natural scenes, they eventually pay more attention to objects than to their contexts, and so as objects and their natural contexts co-occur infants may also learn more about objects as well as objects' typical contexts. Context defines characteristics of the setting in which an object is processed and encoded. Some attributes of object context may be integral to object processing, so context may influence (or give meaning) to how infants come to sense and think about objects. Moreover, it is likely that infants generalize object-context relations to novel instances, perhaps along the lines of statistical learning (e.g., Saffran, 2009)

Finally, our results suggest rethinking how images of objects are presented in infancy and perception studies. Infants scanned objects and contexts in natural scenes differently from the same objects in neutral contexts. Presenting infants with objects in their natural contexts eliminates persistent lurking criticisms surrounding the use of artificial or digitally manipulated stimuli (Neisser, 1976; Schmuckler, 2001) and affords a more ecologically valid assessment of infant perception and cognition (Bronfenbrenner, 1979; Gibson, 1979). As von Hofsten (1983, pp. 243–244) opined, "If the infant has any preadapted means for extracting information about the environment, it is reasonable to believe that the information to be extracted is … ecologically valid … for the infant."

Acknowledgments

References

Bartels A, Zeki S. Functional brain mapping during free viewing of natural scenes. Human Brain Mapping. 2004; 21:75–85.10.1002/hbm.10153 [PubMed: 14755595]

Biederman, I. On the semantics of a glance at a scene. In: Kubovy, M.; Pomerantz, J., editors. Perceptual organization. Hillsdale, NJ: Erlbaum; 1981. p. 213-253.

Biederman I. Recognition-by-components: A theory of human image understanding. Psychological Review. 1987; 94:115–117.10.1037/0033-295X.94.2.115 [PubMed: 3575582]

Bornstein MH, Mash C, Arterberry ME. Perception of object-context relations: Eye-movement analysis in infants and adults. Developmental Psychology. (in press).

Bornstein, MH.; Hahn, C-S.; Suwalsky, JTD.; Haynes, OM. Socioeconomic status, parenting, and child development: The *Hollingshead Four-Factor Index of Social Status* and the *Socioeconomic Index of Occupations*. In: Bornstein, MH.; Bradley, RH., editors. Socioeconomic status, parenting, and child development. Mahwah, NJ: Erlbaum; 2003. p. 29-82.

Bronfenbrenner, U. The ecology of human development. Cambridge, MA: Harvard University Press; 1979.

Butler J, Rovee-Collier C. Contextual gating of memory retrieval. Developmental Psychology. 1989; 22:533–552.10.1002/dev.420220602

Colombo J, Laurie C, Martelli T, Hartig B. Stimulus context and infant orientation discrimination. Journal of Experimental Child Psychology. 1984; 37:576–586.10.1016/0022-0965(84)90077-8 [PubMed: 6747549]

Davenport JL, Potter MC. Scene consistency in object and background perception. Psychological Science. 2004; 15:559–564.10.1111/j.0956-7976.2004.00719.x [PubMed: 15271002]

Gibson, JJ. The ecological approach to visual perception. Boston: Houghton-Mifflin; 1979.

Gredebäck G, Johnson S, von Hofsten C. Eye tracking in infancy research. Developmental Neuropsychology. 2010; 35:1–19.10.1080/87565640903325758 [PubMed: 20390589]

Haaf RA, Lundy BL, Coldren JT. Attention, recognition, and the effects of stimulus context in 6-month-old infants. Infant Behavior and Development. 1996; 19:93–106.10.1016/S0163-6383(96)90047-8

Hollingshead, AB. Unpublished manuscript. Yale University; 1975. The four-factor index of social status.

Li F, VanRullen R, Koch C, Perona P. Rapid natural scene categorization in the near absences of attention. Proceedings of the National Academy of Sciences USA. 2002; 99:9596–9601.10.1073/pnas.092277599

Nadel, L.; Willner, J.; Kurz, EM. Cognitive maps and environmental context. In: Balsam, PD.; Tomie, A., editors. Context and learning. Hillsdale, NJ: Erlbaum; 1985. p. 385-406.

Neisser, U. Cognition and reality: Principles and implications of cognitive psychology. New York: W. H. Freeman and Company; 1976.

Palmer SE. The effects of contextual scenes on the identification of objects. Memory & Cognition. 1975; 3(5):519–526. Retrieved from http://mc.psychonomic-journals.org/.

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin. 1998; 124(3):372–422. Retrieved from http://www.apa.org/pubs/journals/bul/. [PubMed: 9849112]

Rayner K, Smith TJ, Malcolm GL, Henderson JM. Eye movements and visual encoding during scene perception. Psychological Science. 2009; 20:6–10.10.1111/j.1467-9280.2008.02243.x [PubMed: 19037907]

Righart R, de Gelder B. Context influences early perceptual analysis of faces—an electrophysiological study. Cerebral Cortex. 2006; 16:1249–1257.10.1093/cercor/bhj066 [PubMed: 16306325]

Rolls ET, Aggelopoulos NC, Zheng F. The receptive fields of inferior temporal cortex neurons in natural scenes. The Journal of Neuroscience. 2003; 2(1):3, 339–348. Retrieved from http://www.jneurosci.org/.

Rovee-Collier C, DuFault D. Multiple contexts and memory retrieval at three months. Developmental Psychobiology. 1991; 24:39–49.10.1002/dev.420240104 [PubMed: 2015962]

Saffran, JR. What can statistical learning tell us about infant learning. In: Woodward, A.; Needham, A., editors. Learning and the infant mind. New York, NY: Oxford University Press; 2009. p. 29-46.

Schmuckler MA. What is ecological validity? A dimensional analysis. Infancy. 2001; 2:419–436.10.1207/S15327078IN0204_02

von Hofsten C. Foundations for perceptual development. Advances in Infancy Research. 1983; 2:241–264.10.1336/0893911135

Williams A, Weisstein N. Line segments are perceived better in coherent context than alone: An object-line effect in visual perception. Memory & Cognition. 1978; 6(2):85–90. Retrieved from http://mc.psychonomic-journals.org/.

Figure 1.
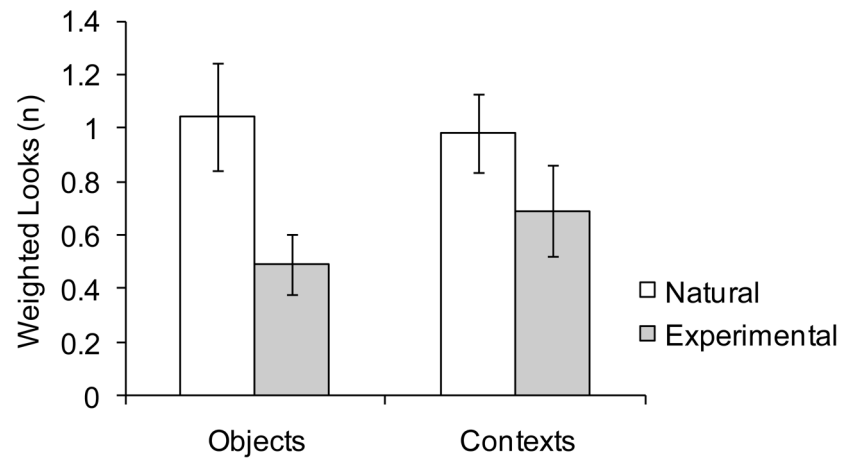Examples of (A) natural scenes and (B) experimental scenes.

Figure 2.
Weighted mean numbers of infant fixations (± SEM) for objects and contexts in natural and experimental scenes.