INFANCY, *12*(1), 31–43

Mechanisms of Visual Object Recognition in Infancy: Five-Month-Olds Generalize Beyond the Interpolation of Familiar Views

Clay Mash
National Institute of Child Health and Human Development
National Institutes of Health, U.S. Department of Health and Human Services

Martha E. Arterberry
Department of Psychology
Gettysburg College

Marc H. Bornstein
National Institute of Child Health and Human Development
National Institutes of Health, U.S. Department of Health and Human Services

This work examined predictions of the interpolation of familiar views (IFV) account of object recognition performance in 5-month-olds. Infants were familiarized to an object either from a single viewpoint or from multiple viewpoints varying in rotation around a single axis. Object recognition was then tested in both conditions with the same object rotated around a novel axis. Infants in the multiple-views condition recognized the object, whereas infants in the single-view condition provided no evidence for recognition. Under the same 2 familiarization conditions, infants in a 2nd experiment treated as novel an object that differed in only 1 component from the familiar object. Infants' object recognition is enhanced by experience with multiple views, even when that experience is around an orthogonal axis of rotation, and infants are sensitive to even subtle shape differences between components of similar objects. In general, infants' performance does not accord with the predictions of the IFV model of object recognition. These findings motivate the extension of future research and theory beyond the limits of strictly interpolative mechanisms.

Send correspondence to Clay Mash, NICHD Section on Child and Family Research, 6705 Rockledge Drive, Suite 8030, Bethesda, MD, 20892. E-mail: mashc@mail.nih.gov

Research on the origins of visual object recognition has revealed that infants as young as 3 months of age can detect and recognize structural properties of three-dimensional objects that are invariant over transformations of spatial orientation (Caron, Caron, & Carlson, 1979; Cook, Field, & Griffiths, 1978; Kellman, 1984; Ruff, 1978). Despite this volume of work on object recognition performance, little effort has been devoted to understanding the internal processes and representational formats that underlie these abilities so early in life.

The central problem for explanations of visual object recognition is accounting for how viewers readily identify objects when the two-dimensional retinal projections of those objects differ from one encounter to another. Several models have been proposed to account for this ability (see Palmeri & Gauthier, 2004; Pinker, 1984, for overview). One approach that has motivated a considerable amount of research is the image-based viewpoint-dependent approach, which proposes that viewers store images of objects in memory, and that recognition is based on relating objects currently in view to the stored images (e.g., Tarr & Pinker, 1989). Although seemingly uneconomical with respect to storage demands, a large number of viewpoint effects on recognition speed and accuracy have been reported (e.g., Tarr, Williams, Hayward, & Gauthier, 1998). Generally, as the orientation of an object increasingly differs from familiar views, recognition performance (e.g., speed and accuracy) correspondingly decreases (see Tarr & Bulthoff, 1995, for review).

A specific processing mechanism that has been proposed to account for generalizations from familiar to novel views is interpolation between familiar views (IFV; Bulthoff & Edelman, 1992). By this account, objects can be recognized in novel orientations to the extent that a matching view can be generated through an interpolation between two views already in storage. Under these circumstances, success depends on a common axis of rotation uniting the stored views with the present view. The IFV hypothesis generates specific predictions about the nature and effects of viewpoint experience in establishing a basis for subsequent object recognition. Such a mechanism accounts best for novel viewpoints that are intermediate between familiar views that all relate to a common axis of rotation. Thus, additional experience with and storage of viewpoints derived from other axes of rotation is not expected to support recognition.

These predictions can be tested experimentally by providing viewpoint experience with a novel object around a single axis of rotation, and examining recognition of the same object when it has been rotated around a novel axis. If infants rely on strictly interpolative mechanisms, they should have difficulty recognizing the item in the novel orientation regardless of whether they have experience with a single, different view or with multiple views around an unrelated axis. Neither type of experience offers a match by interpolation. If, alternatively, infants show a benefit from experience with rotation around an unrelated axis compared to performance based on experience with only a single view, something other than an interpolative mechanism must be involved (cf. Peissig, Wasserman, Young, & Biederman, 2002). This study examined whether infants' recognition performance coincides with the predictions of the IFV hypothesis.

In the first experiment, 5-month-old infants were familiarized to static images of a single, novel object. Over the familiarization phase, half of the infants viewed the same image of the object across the four trials (thus receiving a single view; see Figure 1a). The other half of the infants viewed four different images of the same object with each showing a different view around the same axis of rotation (Figure 1a). Following familiarization, two test trials were presented, one familiar and one novel. Infants viewing a single orientation of the familiarization object were presented with the same object in the same orientation on one trial, and the same object in a novel orientation on the other. Infants who experienced multiple views of the object saw the same object in a novel rotation around the same axis of rotation, and the same object rotated around a novel axis.

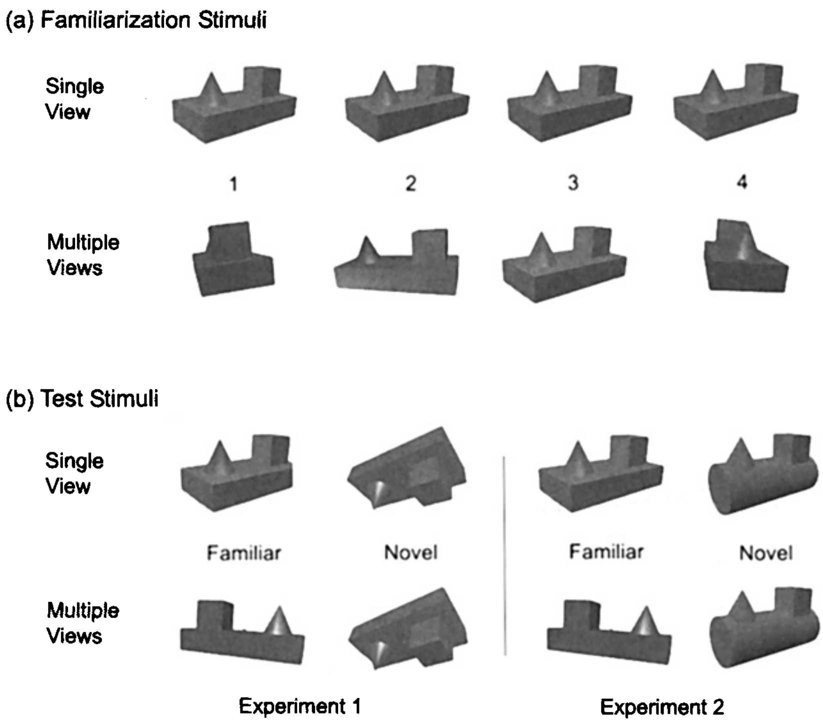

FIGURE 1 (a) Familiarization stimuli used in Experiments 1 and 2. Stimuli and order were randomized in the multiple-views condition. (b) Test stimuli used in Experiments 1 and 2.

Following conventions of the familiarization/novelty preference paradigm (e.g., Bornstein, 1985), a decrement in looking time over the course of trials in the familiarization phase was interpreted to indicate visual encoding and corresponding recognition. A systematic increase in looking at a test stimulus was interpreted as a failure of recognition (i.e., the discrimination of the test item as different from what was seen before). If infants utilize a mechanism based exclusively on the interpolation of familiar views when processing visual objects for recognition, one would expect neither of the two familiarization conditions to enable recognition of the novel test item. Neither condition offers a match of the novel test item via strict interpolation. If infants in either condition do recognize the novel test item, then some other class of processing mechanism must be involved.

EXPERIMENT 1

Methods

Participants. Thirty-eight infants aged 5 months ($M = 22.1$ weeks, $SD = 0.8$) were included in the final sample (17 boys and 21 girls). An additional 4 infants participated but were not included due to fussiness or experimenter error.

Stimuli. A series of six graphic images all representing a single novel object was used (Figure 1). The object consisted of a cube and cone each attached to the top surface of a brick-shaped solid. The object images were rendered with a graphical software package. Five of six distinct views of the object varied only in rotation around a vertical axis and served as the familiarization stimuli. Each of these five views differed from its neighboring views by 60° in rotation. The sixth view was a 120° rotation around the object's horizontal axis from the midpoint of the first five orientations. The height of the object images subtended 11.2° of visual angle, with their width varying between 12.2° and 18.3° across the different rotations. The amount of rotation of the test stimulus was identical between conditions relative to the midpoint of their familiarization sequences. Additionally, the horizontal-axis rotation was equivalent to the average amount of rotation across the vertical-axis views sequence.

Two additional images were used in the experiment. A colorful animation of a carousel centered on a black background was used between each trial to recenter infants' fixation of the stimulus display before the next trial's image was presented. Additionally, an unrelated image (18.5° visual angle in height) that consisted of a red square, yellow circle, and blue triangle against a white background was used to index infants' attentional engagement both at the beginning (pretest) and the end (posttest) of the session.

Procedure. Infants were seated directly in front of the stimulus display at a viewing distance of approximately 36 cm in a dimly lit room. On each trial, the spinning carousel image appeared first, and the experimenter observed the infant's fixation in the video monitor. The experimenter initiated the presentation of each trial after infants were judged to be fixating the carousel. Infants' fixation of each image was coded using a corneal reflection technique. The experimenter depressed a timing key when the reflected image was superimposed over the pupil, and released the key when the reflection moved off the pupil with each change in fixation.

Infants were randomly assigned to one of two experimental conditions: single view or multiple views. A fixed-trials procedure was used: For every infant, eight 10-sec trials were administered. On the first trial, the pretest–posttest image was presented. Then, four familiarization trials were presented, being constituted differently between the two familiarization conditions. For infants in the multiple-views condition, four randomly selected vertical-axis rotations were presented in an order uniquely randomized for each infant. Infants in the single-view condition were presented only with the midpoint rotation stimulus over the four trials. In both conditions, two test trials were then administered in a fixed order, the first being a familiar test trial and the second being a novel test trial. For infants in the single-view condition, the familiar test stimulus was the same that they had seen during familiarization. For infants in the multiple-views condition, the familiar test stimulus was the remaining member of the vertical-axis rotation series that they had not seen during familiarization. For infants in both conditions, the novel test stimulus was the horizontal-axis rotation. Following the test trials, the pretest– posttest stimulus was presented again for 10 sec, completing the session. The intertrial interval varied somewhat according to the time required for the infant to reestablish fixation, but generally did not exceed 3 sec.

To examine coding reliability of the primary experimenter, a second experimenter coded infants' looking time from videotape for 20% of the sessions. Over individual trials and sessions, $r = .94$.

Results and Discussion

Preliminary analyses revealed no effects of gender, $F(1, 34) = 0.33$, *ns,* so subsequent analyses collapsed across boys and girls. To examine looking time during the familiarization phase, averages were taken of the first two (baseline) and last two (end of familiarization) trials. A mixed-model analysis of variance (ANOVA) was used to compare looking times between viewpoint conditions as a between-subjects factor and phases of the trial sequence as a within-subjects factor. The trial sequence phase factor was examined using five orthogonal contrasts: baseline versus end of familiarization, end of familiarization versus familiar test, familiar test versus novel test, novel test versus posttest, and posttest versus pretest. Effect sizes were calculated as partial eta squared (η_p^2).

The analysis revealed a significant difference between baseline and the end of familiarization, $F(1, 36) = 4.50$, $p = .041$, $\eta_p^2 = .11$, with infants looking longer early than later (see Figure 2). This contrast did not interact with familiarization condition. Additionally, the analysis revealed a significant interaction between the familiar test versus novel test contrast and familiarization condition, $F(1, 36) = 9.79$, $p = .003$, $\eta_p^2 = .21$. Tukey post-hoc tests of simple effects revealed that infants in the single-view condition looked significantly longer at the novel test stimulus than the familiar test stimulus, $t(18) = 3.00$, $p = .008$, whereas infants in the multiple-views condition did not, $t(18) = 1.43$, *ns*. No other effects were observed. Importantly, there was no evidence that general fatigue may have affected performance either in the familiarization or the test trials in either condition (contrast between pretest and posttest, $F(1, 36) = 1.03$, *ns;* interaction with familiarization condition, $F(1, 36) = .07$, *ns*). Finally, there was also a significant difference between the novel test and posttest, $F(1, 36) = 12.54$, $p = .001$, $\eta_p^2 = .26$, with infants looking longer at the posttest stimulus.

Infants having experience with multiple viewpoints of a novel object around one axis generalized looking time to the same object when it was rotated around a completely different axis. Infants having experience with a singular view, however, did not provide evidence of recognizing the stimulus object when it appeared in a novel orientation at test. Infants' object recognition performance appears to have benefited from experience with multiple viewpoints, enabling recognition even when the test item's orientation novelty derived from a different rotation axis relative to the experience obtained during familiarization. This pattern of findings is not predicted by the IFV hypothesis.

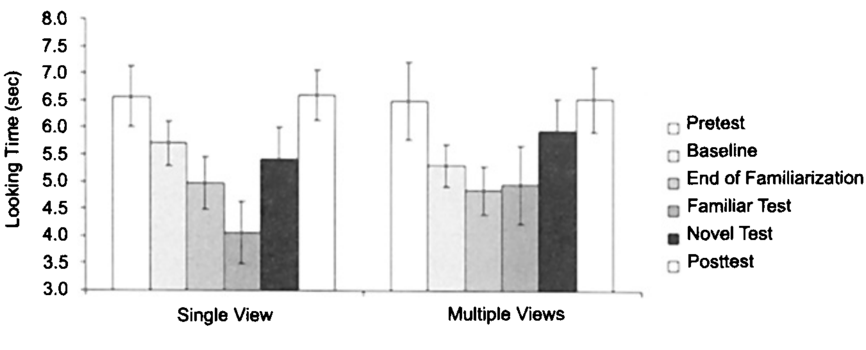

FIGURE 2 Mean looking time by condition and trial phase for Experiment 1. Error bars represent 1 *SEM*.

Before firm conclusions can be drawn, additional alternatives must also be considered. One possibility is that infants in the single-view condition actually did recognize the test item (novel rotation) as the same object seen during familiarization, yet nevertheless looked longer because they found the new orientation novel. Although infants in the single-view condition may have dishabituated more to axis novelty than to object novelty per se, they clearly did not treat the familiar and novel test items as equivalent stimuli. Another possibility may be that, whereas infants in the multiple-views condition formed a three-dimensional representation of the stimulus object during familiarization, infants in the single-view condition formed only a two-dimensional representation of the image that was then compared to the test item simply on the basis of contour overlap. This explanation, however, does not hold merit in light of previous findings indicating that infants the same age do generalize familiarity of multipart two-dimensional stimuli across changes in orientation (Schwartz & Day, 1979), and furthermore perceive equivalence between objects and their two-dimensional images (DeLoache, Strauss, & Maynard, 1979). Such findings support the interpretation here that infants in the single-view case encoded a novel object and failed to recognize that object when it was seen in a novel orientation.

An additional aspect of the results that warrants further consideration is the basis on which infants in the multiple-views condition generalized looking time between the two test stimuli. Did infants in the multiple-views condition actually recognize the novel test item? One alternative account of this finding is that the stimulus variability present in the multiple-views condition may have rendered the familiarization set more difficult to encode than that of the single-view condition. If so, infants in the former condition may have formulated a less discriminating representation simply due to incomplete learning during familiarization. This alternative does not seem plausible, however, given the equivalence between conditions in looking time decrement that was observed during the familiarization phase (the interaction between familiarization condition and the contrast between early familiarization and end of familiarization was not significant, $F(1, 36) = 0.08$, *ns*). Such uniform and statistically equivalent habituation is most readily interpreted as equivalent learning between conditions.

Further consideration of the results raises an additional question about whether infants in the multiple-views condition generalized their looking time during the test phase not because they identified the test object as the same as what they had seen before, but because their difference detection tolerance was increased by the constant variability within the familiarization set. This also seems unlikely because infants in both familiarization conditions clearly dishabituated to the posttest stimulus, looking longer at it than the novel test stimulus, and thus were still engaged in the task and making systematic distinctions between stimuli throughout the task. Even though the present findings are inconsistent with an account based on a drastic reduction in visual discrimination, it remains possible that the familiarization phase in the multiple-views condition increased behavioral generalization for a range of relatively similar objects. If so, the generalization of looking time to the novel test stimulus in that condition does not provide clear evidence for recognition.

Finally, one more question arising from the performance observed in Experiment 1 is whether infants might require multiple views of an object to form a discriminating representation, or if single-view experience enables the discrimination of similar but distinct objects. To examine both issues in relation to the stimuli used in Experiment 1, a second experiment was conducted. Infants were familiarized either with a single view or with multiple views of the same object used previously, and then tested with a similar object differing only in the shape of one component.

EXPERIMENT 2

Methods

Participants. Thirty-six 5-month-olds (M = 22.3 weeks, SD = 2.0) were included in the final sample (21 boys and 15 girls). An additional 3 infants participated but were not included due to fussiness.

Stimuli and procedure. The same familiarization stimuli, apparatus, and procedures used in Experiment 1 were used. A new test stimulus was constructed for each condition by replacing the base brick with a cylinder (see Figure 1b). Interrater reliability was examined in the same manner described earlier (r = .91).

To enable the comparison of results between experiments, a comparison of baseline preference was conducted between the two test stimuli: the brick-based stimulus used in Experiment 1 and the cylinder-based stimulus used in this experiment. Each stimulus was presented to a separate group of 20 infants (M age = 22.4 weeks, SD = 2.8; 10 boys and 10 girls) in alternating order over the course of six 10-sec trials using the same apparatus and procedures used in Experiment 1. The starting stimulus was counterbalanced across participants, and fixation time was measured on each trial. There was no significant difference between the two test stimuli in the amount of time each was fixated in the absence of prior familiarization, $t(19) = 0.20$, *ns*.

Results and Discussion

The looking times throughout the familiarization phase were averaged over two blocks, one representing the first two trials (baseline) and the other representing the last two trials (end of familiarization). A mixed-model ANOVA was used to compare looking times between viewpoint conditions as a between-subjects factor and phases of the trial sequence as a within-subjects factor. The trial sequence phase factor was examined using five orthogonal contrasts: baseline versus end of familiarization, end of familiarization versus familiar test, familiar test versus novel test, novel test versus posttest, and posttest versus pretest. The trial phase factor was comprised of planned comparisons representing effects of habituation (baseline vs. end of familiarization) and recognition (end of familiarization vs. test). The analysis revealed a significant difference between familiar test and novel test, $F(1, 34) = 10.05$, $p = .003$, $\eta_p^2 = .23$, with infants looking longer at the novel test stimulus (see Figure 3). This contrast did not interact with familiarization condition, $F(1, 34) = 0.24$, *ns*, revealing an equivalent novelty preference between groups. No other effects were significant.

Unlike the performance observed in Experiment 1, infants who experienced many different viewpoints of one object did not generalize looking to the novel test. In this case, the novel test item was a novel object, yet was also very similar to the familiarization object, differing in only one component. Despite the relative subtlety of this difference, infants detected it and discriminated the novel test object from the object they had seen during the familiarization phase. These findings appear to rule out the possibility that infants in the multiple-views condition of Experiment 1 generalized looking at the test stimulus merely on the basis of a transient elevation of their difference detection threshold. Instead, the more likely account of infants' performance in Experiment 1 is that they recognized the object in a novel rotation as the same one that they had seen during familiarization. Additionally, with infants in both familiarization conditions discriminating the test stimuli, these findings also indicate that 5-month-olds do not need to have multiple-views experience with novel objects to discriminate them from others; a single view is adequate even when the objects are relatively similar.

Additional analyses were conducted to compare test trial performance between Experiments 1 and 2. The first analysis examined data from the single-view con-

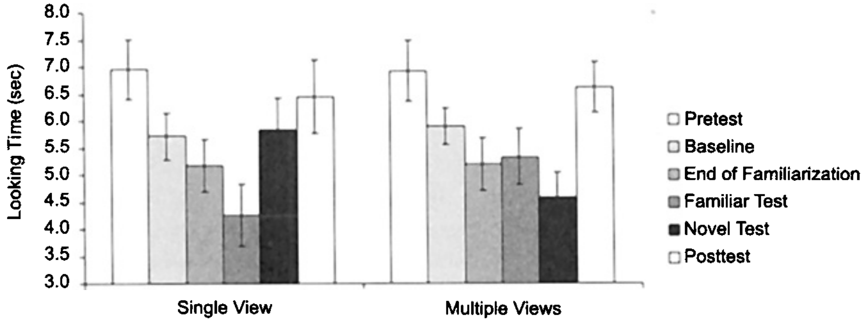

FIGURE 3 Mean looking time by condition and trial phase for Experiment 2. Error bars represent 1 *SEM*.

dition in both experiments, and included test trial (familiar vs. novel) as a within-subjects factor and test stimulus (brick-based vs. cylinder-based) as a between-subject factor. The analysis revealed a significant difference between test trials, $F(1, 35) = 15.77$, $p < .001$, $\eta p^2 = .31$, with infants looking longer at the novel test stimuli. No other effects were significant. Thus infants treated as equally novel the novel rotation seen in Experiment 1 and the novel object seen in Experiment 2.

The second analysis examined data from the multiple-views condition in both experiments, and included test trial (familiar vs. novel) as a within-subjects factor, and test stimulus (brick-based vs. cylinder-based) as a between-subject factor. The analysis revealed a significant interaction between test trial and test stimulus, $F(1, 35) = 5.45$, $p = .025$, $\eta_p^2 = .14$. Tukey post-hoc tests reveal that infants viewing the cylinder-based novel test item (Experiment 2) fixated it longer than those viewing the novel rotation of the familiar brick-based test item (Experiment 1), $t(17) = 2.52$, $p = .022$. This finding confirms the interpretation that with multiple viewpoint experience, 5-month-olds' visual recognition is both flexibly general to orientation variation and simultaneously specific to even subtle differences between objects.

GENERAL DISCUSSION

The results of these experiments reveal that multiple viewpoint experience enhances 5-month-old infants' recognition of objects, and that recognition generalization is not restricted to a unified common axis of orientation variability. Additionally, when orientation is held relatively constant between different objects, the specificity of infants' object representations is high, enabling discrimination of subtly different objects on the basis of experience with single or multiple views. Thus, despite their capacity to formulate precise representations, infants' representation of multiple viewpoints enables relatively broad orientation generalization in recognition.

With respect to processing mechanisms that can conceivably underlie these abilities, 5-month-old infants' performance is not best characterized by a mechanism that interpolates between familiar views. Such an approach predicts recognition failure when there is no range of viewpoints between which to interpolate for a match to memory (i.e., the single-view condition examined presently), or when the test item's orientation varies substantially from possible interpolations of previous views (i.e., the multiple-views condition). Yet performance under these competing viewpoint conditions was clearly not equivalent, with infants failing only in the single-views condition. Results of this sort require a different explanation.

In Experiment 1, recognition appears to have depended on a constructive process (i.e., infants recognized the novel test item only after serial exposure to different views). Unlike what the IFV view predicts, however, viewpoint generalization extended beyond a single axis of rotation. This observation suggests that object-centered structural invariants were extracted over different views, stored in memory, and used in recognition matching to the test items. As noted initially, previous work suggests that infants are capable of detecting structural invariants of specific objects, and representing them internally (e.g., Caron et al., 1979; Cook et al., 1978). Slater, Mattock, Brown, and Bremner (1991) reported facilitative effects of multiple-orientation experience even in newborns' recognition of simple two-dimensional line patterns within days of birth. The extraction and storage of structural invariants appears to be a fundamental aspect of visual recognition, perhaps undergoing relatively little change despite other developments in visual memory (e.g., Olson & Strauss, 1984). The work reported here extends such findings by addressing the range of possible processing mechanisms that might underlie this ability.

The findings reported here are not consistent with those of an earlier study that utilized comparable methods. Kellman (1984) familiarized 4-month-old infants to images of an object in rotation around two different axes under either moving or static conditions, and then tested for discrimination of a novel object. Infants seeing static images during familiarization failed to discriminate the test items, but infants seeing moving objects did discriminate. These findings suggest that object motion may be necessary for infants to extract three-dimensional shape (see also Arterberry & Yonas, 1988, 2000), yet infants in the study reported here did so under static conditions. There are several factors that might account for this salient difference between studies. For one, infants in this study were 1 month older than those examined by Kellman (1984), and therefore may have simply utilized more mature perceptual mechanisms.

Another potentially relevant difference is that Kellman's (1984) familiar–novel test comparison may have been less discriminable than ours. Whereas his stimuli were comprised entirely of linear edges, the stimuli in these experiments differed one from the other in edge curvature. Infants may process straight and curved edges differently in visual equivalence judgments (Fantz, Fagan, & Miranda, 1975; Van de Walle, 1997), perhaps enhancing the discriminability of our test comparison relative to Kellman's. Finally, Kellman (1984; see also Kellman & Short, 1987) habituated infants to criterion, and did so under relatively lengthy conditions involving 6 to 8 sec between trials, whereas we administered a fixed and relatively small number of trials in brief succession. Thus, Kellman's procedure may have reduced infants' attention to the static test stimuli via fatigue, whereas the task reported here clearly did not. In contrast to Kellman's (1984) conclusions, the findings of these experiments reveal that infants appear capable of extracting three-dimensional form from static images under some conditions.

A fundamental question that remains is what kind of storage and processing mechanism could better account for infants' performance. One possibility may be the structural description approach. Proponents of this view have argued that visually attended objects are decomposed into primitive elements that are then matched to elemental object descriptions stored in memory (e.g., Biederman, 1987; Marr & Nishihara, 1978). Representations of this kind support recognition independently of the viewpoint range that is experienced visually as long as an adequate sampling of visual structure is obtained (see Haaf et al., 2003, and Mash, 2006, for related discussion about visual development). Recently, Stankiewicz (2003) pointed out that as the contrasting theoretical approaches have evolved, viewpoint effects *and* viewpoint-independent performance are both predicted by models from both predominant approaches (view based and structural description). Thus, viewpoint effects like those reported here may ultimately constitute only a partial basis for distinguishing general theories of object recognition. More specifically, however, this work does enable the elimination of a longstanding hypothesis whose implications for development have been unknown. These findings motivate search beyond strictly interpolative mechanisms to explain visual object recognition near the beginning of life.

ACKNOWLEDGMENTS

Martha Arterberry is now at the Department of Psychology, Colby College, Waterville, ME. This work was supported by the Intramural Research Program of the National Institute of Child Health and Human Development, National Institutes of Health. We thank Nan Cummins, Debra Falk, Sarah Jones, Erica Moran, Elizabeth Seiver, and Erin Stine for assistance with data collection and scoring. We thank Peter Gerhardstein and the anonymous reviewers for comments on an earlier version of this article.

REFERENCES

Arterberry, M. E., & Yonas, A. (1988). Infants' sensitivity to kinetic information for three-dimensional object shape. *Perception and Psychophysics, 44,* 1–6.

Arterberry, M. E., & Yonas, A. (2000). Perception of structure from motion by 8-week-old infants. *Perception and Psychophysics, 62,* 550–556.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Bornstein, M. H. (1985). Habituation of attention as a measure of visual information processing in human infants: Summary, systematization, and synthesis. In G. Gottlieb & N. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 253–300). Westport, CT: Ablex.

Bulthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences, 89,* 60–64.

Caron, A. J., Caron, R. F., & Carlson, V. R. (1979). Infant perception of the invariant shape of objects varying in slant. *Child Development, 50,* 716–721.

Cook, M., Field, J., & Griffiths, K. (1978). The perception of solid form in early infancy. *Child Development, 49,* 866–869.

DeLoache, J. S., Strauss, M. S., & Maynard, J. (1979). Picture perception in infancy. *Infant Behavior and Development, 2,* 77–89.

Fantz, R. L., Fagan, J. F., & Miranda, S. B. (1975). Early visual selectivity as a function of pattern variables, previous exposure, age from birth and conception, and expected cognitive deficit. In L. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. 1, pp. 249–345). New York: Academic.

Haaf, R. A., Fulkerson, A. L., Jablonski, B. J., Hupp, J. M., Schull, S. S., & Pescara-Kovach, L. (2003). Object recognition and attention to object components by preschool children and 4-month-old infants. *Journal of Experimental Child Psychology, 86,* 108–123.

Kellman, P. J. (1984). Perception of three-dimensional form by human infants. *Perception and Psychophysics, 36,* 353–358.

Kellman, P. J., & Short, K. R. (1987). Development of three-dimensional form perception. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 545–557.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B: Biological Sciences, 200,* 269–294.

Mash, C. (2006). Multidimensional shape similarity in the development of visual object classification. *Journal of Experimental Child Psychology, 95,* 128–152.

Olson, G. M., & Strauss, M. S. (1984). The development of infant memory. In M. Moskovitch (Ed.), *Infant memory: Its relation to normal and pathological memory in humans and other animals* (pp. 29–48). New York: Plenum.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Neuroscience Reviews, 5,* 1–13.

Peissig, J. J., Wasserman, E. A., Young, M. E., & Biederman, I. (2002). Learning an object from multiple views enhances its recognition in an orthogonal rotational axis in pigeons. *Vision Research, 42,* 2051–2062.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18,* 1–63.

Ruff, H. A. (1978). Infant recognition of the invariant form of objects. *Child Development, 49,* 293–306.

Schwartz, M. S., & Day, R. H. (1979). Visual shape perception in early infancy. *Monographs of the Society for Research in Child Development, 44*(Whole No. 7).

Slater, A., Mattock, A., Brown, E., & Bremner, J. G. (1991). Form perception at birth: Cohen and Younger (1984) revisited. *Journal of Experimental Child Psychology, 51,* 395–406.

Stankiewicz, B. J. (2003). Just another view. *Trends in Cognitive Sciences, 7,* 526.

Tarr, M. J., & Bulthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1494–1505.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233–282.

Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience, 1,* 275–277.

Van de Walle, G. A. (1997). *Perceptual foundations of categorization in infancy: The animal/vehicle distinction.* Unpublished doctoral dissertation, Cornell University, Ithaca, NY.