

Anticipatory Eye Movements Reveal Infants' Auditory and Visual Categories

Bob McMurray and Richard N. Aslin
Department of Brain and Cognitive Sciences
University of Rochester

We introduce a new paradigm for the assessment of auditory and visual categories in 6-month-old infants using a 2-alternative anticipatory eye-movement response. Infants were trained by 2 different methods to anticipate the location of a visual reinforcer at 1 of 2 spatial locations (right or left) based on the identity of 2 cuing stimuli. After a training phase, infants were presented with a series of generalization trials in which novel (untrained) stimuli served as the cue to the anticipatory eye movement. Four experiments illustrated that infants can learn the 2-choice discriminative response during training. Infants also showed anticipatory eye movements to novel stimuli, indicating sensitivity to variations along a variety of stimulus dimensions (e.g., color, shape, orientation, spatial frequency, pitch, and duration). In addition, the paradigm can be used to assess categorization in individual infants, thereby revealing the stimulus dimensions to which infants naturally attend.

One of the most important cognitive skills acquired during infancy is categorization, the ability to group together on some basis (e.g., perceptual or functional) a set of exemplars that are easily discriminated from each other. The study of infants' categorization abilities has a long and rich history (see Cohen, 2000; Mandler, 1998). However, one of the most vexing problems in this literature is the absence of a methodology that provides a definitive measure of categorization. In older children and adults, such a measure consists of a verbal response or a nonverbal behavior (e.g., pointing or button-pressing) elicited by a verbal query. In preverbal infants, inferences about categorization have relied on methods based on habituation, conjugate reinforcement, and operant conditioning.

Supplementary materials to this article are available on the World Wide Web at <http://www.infancyarchives.com>

Requests for reprints should be sent to Bob McMurray, Department of Psychology E11 SSH, University of Iowa, Iowa City, IA 52240. E-mail: bob-mcmurray@uiowa.edu

In the *multiple-exemplar-habituation* (MEH) paradigm, infants are presented with two or more discriminably different exemplars, in random order, and their duration of looking is measured across a series of trials until it reaches a predetermined decrement from an initial baseline. During the posthabituation test phase, if infants fail to exhibit recovery of looking time to a novel token that shares the criterial attribute with the familiarized exemplars, categorization is inferred. In contrast, if infants cannot extract the invariant attribute across multiple exemplars, or they have memorized each exemplar, then they should discriminate (i.e., show recovery of looking time) to this novel within-category token (for a classic example of this paradigm, see Fagan, 1976).

Conjugate reinforcement procedures have relied on high-amplitude sucking (HAS) and foot-kicking. As in MEH two or more exemplars can be presented during the learning or habituation phase in HAS (see Juszyk & Derrah, 1987) or during the acquisition phase in foot-kicking (see Hayne, 1996). A decline in foot-kicking or continued habituation in HAS in the presence of a novel, discriminable stimulus suggests that this stimulus is a member of the same category as the training stimuli.

Finally, operant conditioning has been used to train infants to make head movements in response to a particular stimulus. Kuhl (1979, 1983) used this *conditioned head-turning* paradigm to assess categorization by beginning the task with a simple discrimination (a single repeating background token and a single target token) and then adding multiple tokens to these background and target sets. Infants were visually reinforced for head turns to the target set but not for head turns to the background set. Although they initially made head turns (false alarms) to these discriminably different background tokens, they quickly learned to suppress these head turns and to only turn to the target set. Thus, infants as young as 6 months of age can learn to attend to certain criterial attributes of speech stimuli and select those that lead to a visual reward, while ignoring other task-irrelevant stimulus variations.

The fundamental problem with both the MEH and conditioning paradigms as a measure of categorization is that they rely on a single response. That is, infants can indicate (by recovery of looking time or by response generalization) that they treat a particular test stimulus as same or different, and by inference as outside or inside the category. However, we know from a century of psychophysical studies that single-response (or yes–no) paradigms are subject to criterion effects. That is, the willingness of the observer to judge two stimuli as different can vary considerably, thereby rendering a decision about the threshold for category membership uncertain. For example, in deciding that a novel exemplar is in the same category as a given exemplar (e.g., a particular dog), participants could base their decision on whether the two exemplars are the same breed of dog (e.g., collie), are both dogs rather than some other four-legged mammal (e.g., cats), or are both animals.

MEH paradigms (as well as Kuhl's version of the conditioned head-turning procedure) attempt to circumvent this problem by exposing the infant to a number of

different exemplars (e.g., different dogs from within the same breed). This not only establishes a context for basing decisions about category membership, but also exposes the infant to the distribution of stimulus variation within the intended category. This exposure may mask any natural biases infants bring to the task of categorization. For example, Kuhl (1983) trained infants with a variety of vowels that varied on irrelevant dimensions (pitch and talker) and found that they were capable of generalizing vowel categories across such dimensions. However, in an experiment in which infants were trained with only one token, generalization suffered, and infants did not successfully identify all of the novel tokens.

In contrast, two-alternative forced-choice (2AFC) paradigms are criterion-free (chance is always 50%) because the observer must actively choose to make either one response or the other to each stimulus. Unfortunately, only rarely have conditioning techniques using two different responses been adapted successfully to assess infant categorization.¹ Attempts to modify the conditioned head-turning technique to assess auditory categories have been largely unsuccessful because infants do not readily learn to make bidirectional head-turning responses (Aslin & Pisoni, 1980).² However, Husain and Cohen (1981) reported some success in training 10-month-olds to make a bidirectional head-turn response to two visual training stimuli and to generalize their responses to novel exemplars.

Our goal in this series of experiments was to expand the range of paradigms available for studying infant categorization by developing a new 2AFC task. We melded a number of useful properties from several other paradigms and dependent responses. First, to obtain a behavioral response that is learned quickly and maintained over many trials, we capitalized on the ability of infants to anticipate the appearance of a visual stimulus and to be engaged (reinforced) by repeated disappearance and reappearance events (Canfield, Smith, Brezsnyak, & Snow, 1997; Haith, Wentworth, & Canfield, 1993). Second, unlike dependent measures based on head turning, eye movements are very fast, metabolically efficient, and easy to detect using an automated eye-tracking system. Third, we trained infants using two cuing stimuli to make two different criterion responses: eye movements to the left or right of a visual display. Finally, after being trained with these two cuing stimuli, infants were presented with novel test stimuli during a generalization phase to assess whether a particular test stimulus was more like one or the other of the cuing stimuli.

¹However, see DeCasper and Spence (1986) for a two-interval sucking paradigm in which newborns begin their sucking burst during one of two tone intervals to trigger an auditory reinforcer.

²We surmise that this difficulty is due to the relatively high metabolic cost of making a head turn, coupled with a relatively less developed capacity for the volitional control of head movements. In addition, to create a situation that promotes detectable head movements by an observer, the visual reinforcer that elicits a head turn is typically displaced by 45° or more, thereby requiring a large shift in spatial attention.

We believe that a particular advantage of our anticipatory eye movement (AEM) paradigm is that infants are not trained to selectively attend to (or ignore) certain criterial attributes or dimensions. For example, in the MEH paradigm, the experimenter defines for the infant which distribution of stimuli “go together” during the habituation phase. However, rarely do such experiments include a control condition in which the set of stimuli do not conform to a category of interest, thereby assessing whether infants can learn an arbitrary grouping of exemplars (for an exception, see Fodor, Garrett, & Brill, 1975).

Thus, there are two key differences between the AEM paradigm described in this article and those discussed earlier: (a) the ability to provide two responses, setting a task-defined criterion for the categorical response; and (b) a focus on what categories the infant brings to the lab, rather than what categories can be induced in the lab by selective exposure to a set of stimuli. We now describe four experiments demonstrating the feasibility of the AEM paradigm in several domains.

EXPERIMENT 1: SINGLE-DIMENSION CATEGORIZATION WITH ORTHOGONAL-DIMENSION VARIATION

The goal of the first experiment was to use the AEM paradigm to explore how a category defined along one dimension (shape) is influenced by variations along an orthogonal dimension (color or orientation). A number of studies have examined how variation along a single criterial dimension clusters to form categories (e.g., orientation [Bomba, 1984], color [Bornstein, Kessen, & Weiskopf, 1976], and shape [Quinn, Slater, Brown, & Hayes, 2001]). However, it is not clear how variation along independent dimensions (of the category) may affect categorization.

Catherwood, Crassini, and Freiberg (1989) approached this question by habituating 5-month-old infants to a single shape with a particular color and then testing for dishabituation to stimuli that differed in shape, color, or both. Infants dishabituated to changes in color but not to changes in shape. They concluded that the color dimension dominates the shape dimension when infants have a “choice” about which dimension to use during the encoding of a single visual stimulus. However, because only a single stimulus was used during habituation, it is not clear whether infants could use shape as a relevant dimension if given information about its importance. For example, a set of stimuli of the same shape that varied in color could be presented using the MEH paradigm, and dishabituation to a different shape of a familiar color would indicate that infants can attend to shape when color is highly variable.

Likewise, in a series of experiments, Colombo, McCollam, Coldren, Mitchell, and Rash (1990) habituated 10-month-old infants to variations in shape and tested their ability to generalize over color. When color was constant during familiariza-

tion but varied during test (Experiment 3), they found significant biases to generalize by color and not by shape (replicating Catherwood et al.'s, 1989, results). However, when color was varied during familiarization, a shape bias emerged. These results raise questions about the role of training in MEH tasks and also suggest that a two-alternative task may help resolve these issues by allowing the experimenter to determine the type of category formed (without instructing the infant on the relevant variation).

Experiment 1 asked a similar question. Infants were presented with two shapes of the same color during 2AFC training and were then presented with variations in either the color or the orientation of these shapes during a generalization phase. Thus, color and orientation were the orthogonal (noncritical) dimensions, but neither was varied (as they would be in the MEH paradigm) during the training phase. In this way we could determine whether color or orientation was relevant to the encoding of the two shape stimuli during training.

Participants

Participants consisted of 3 male and 7 female infants between 5 and 7 months of age ($M = 28.3$ weeks, $SD = 1.93$). Seven additional infants did not complete the experiment due to fussiness or an inability to establish an eye track. In accordance with American Psychological Association guidelines and university human participant protocols, a list of eligible infants was obtained from the University of Rochester Medical Center birth records. Letters were sent (and follow-up calls made) to all parents on this list to recruit the participants. Parents were paid \$5 for their participation or they received an infant-sized T-shirt for their son or daughter.

Automated Eye Tracking and Calibration

All experiments made use of an ASL Pan/Tilt Model 504 remote eye tracker to assess eye movements. The eye tracker consists of a small infrared camera that captures an image of the eye and locates the pupil and corneal reflection. This camera is mounted on servomotors that are controlled by both an optical tracking algorithm and an external magnetic head tracker (a Polhemus FASTRAC operating at 120 Hz) that locates the eye when optical tracking fails (see Aslin & McMurray, 2004/*this issue*). Given the large number of head movements that infants make (looking around the room, behind them at their mother, or simply reorienting to the screen), the magnetic head tracker has proved crucial in making the eye tracker easy to use by automating the process of finding the eye after it moves out of the field of view of the camera (which would otherwise have to be done manually).

After an infant is seated in the parent's lap, his or her attention is directed to the display screen by presenting a colorful contracting and expanding (looming) circle that can be moved by the experimenter to any of nine points on the screen. The ex-

perimenter manually locates the infant's eye with the eye-tracker camera and calibrates the head tracker (which stores the distance, in camera centered coordinates, between the eye and the head-tracker sensor). At this point the camera is turned over to automatic control, while the experimenter finds the correct threshold values for the detection of the corneal reflection and the pupil. Next the two-point calibration scheme (discussed in Aslin & McMurray, 2004/*this issue*) is used by simply moving the visual stimulus to the top left of the screen, recording eye position as the infant fixates this calibration stimulus, and then repeating this process after the stimulus is moved to the bottom right of the screen. Generally, the entire calibration process can be completed in under 1 min. When calibration proves difficult, it can be skipped and reattempted prior to testing.

Procedure

The training phase consisted of a single block of 30 trials. Each trial began with a cuing stimulus that loomed from a point at the center of the screen to a diameter of approximately 6.5° and shrank back to a point over a period of 2 sec. The cuing stimuli consisted of either a yellow square or a yellow cross, matched for total area and diameter. Coincident with the appearance of the cuing stimulus, one of five random sounds was presented to orient the infant to the screen.³ After the cuing stimulus disappeared there was a delay period followed by 1 of 16 different visual reinforcers presented 11° to the right side (for crosses) or left side (for squares) of the screen. During training, the delay between the offset of the cuing stimulus and the onset of the reinforcer grew from 0 msec to 1,800 msec. The duration of each reinforcer was 2 sec and consisted of repetitively moving animals and shapes that covered 9.8° of the display. The use of 16 different reinforcers (randomly selected) was designed to prevent infants from forming an association between the auditory stimulus and the particular visual characteristics of the reinforcer. A second random sound was played at the onset of the reinforcement animation to maintain the infant's interest. Larger animations were presented in the center of the screen about every five trials to reengage the infant.

After the training phase, the infant was given a short break and then testing began. Testing was identical to training except that the delay between the cuing stimulus and the reinforcer was held constant at 1,800 msec, and the reengaging stimulus was not used. Infants were tested until they became too fussy to continue, yielding an average of 22.2 testing trials. In addition to repetitions of the two cuing stimuli during testing, novel (unreinforced) stimuli consisted of the same square

³Note that this version of the AEM paradigm has also been used to examine auditory categorization by replacing the visual cuing stimulus with a neutral shape and using the auditory stimulus as the independent variable (e.g., McMurray, Spivey, & Aslin, 2000).

and cross with either a change in color (from yellow to orange or red) or orientation (clockwise rotation of 10° and 20°), or both.

Data were coded by trained observers from a videotape containing an image of the participant's head and face and, in a smaller window of the video frame, the image of the scene the infant was viewing with gaze position crosshairs superimposed by the ASL eye tracker. When possible, observers coded gaze position from the crosshair output of the eye tracker, using the larger image of the infant's eyes and head as a backup when the eye tracker failed to provide stable crosshairs. This resulted in 78% of the eye movements being coded from the eye tracker output and the remaining 22% from the larger image of the eyes and head. Data were coded from the offset of the visual stimulus (2,000 msec) until either the onset of the reinforcer at 3,800 msec (for original training trials) or until the onset of the next trial (for generalization trials) at 4,800 msec. Each eye movement was coded as a look toward the right, left, center, or off the display screen.

Results

On average, 22.2 trials were obtained from each infant during testing before they became fussy and the experimental session was terminated. Because the criterial dimension during training was shape (square vs. cross), test trials on which infants anticipated the reinforcer on the side predicted by the shape (despite variations in color and orientation) were scored as correct. A correct anticipation occurred when the infant looked longer toward the correct side of the screen during the coding period; it was scored incorrect if they looked longer on the opposite side. Across infants, for test trials with the originally trained cuing stimuli (the yellow square and the yellow cross), infants were correct on 68.7% of the trials, significantly greater than chance, $t(9) = 4.331$, $p = .002$. Only 1 infant scored below chance.

For the 9 infants who scored above chance on the training stimuli, performance on the generalization stimuli was analyzed (i.e., squares and crosses that were red and orange, rotated by 10° and 20°, or both). The relatively small number of test trials prevented us from conducting a $2 \times 3 \times 3$ analysis of variance (ANOVA; Shape \times Color \times Orientation) to analyze the data because of missing data or because no eye movement was made on some trials.⁴ Thus, a logistic regression was used to provide an overall characterization of the pattern of results during testing. This analysis yielded a significant effect of shape: Infants performed better on squares than crosses, $\text{Wald}(1) = 5.70$, $p = .017$. There was also a significant effect of orientation, $\text{Wald}(1) = 4.62$, $p = .032$, with participants performing best on the

⁴This was in part due to the fact that we used a crossed design where infants were tested on orange 10° squares and red 20° crosses, for example, rather than just testing on color and orientation separately. Although this design does permit the analysis of interactions, it requires 18 testing trials for just one replication. This was more than it was possible to obtain reliably across all of the infants.

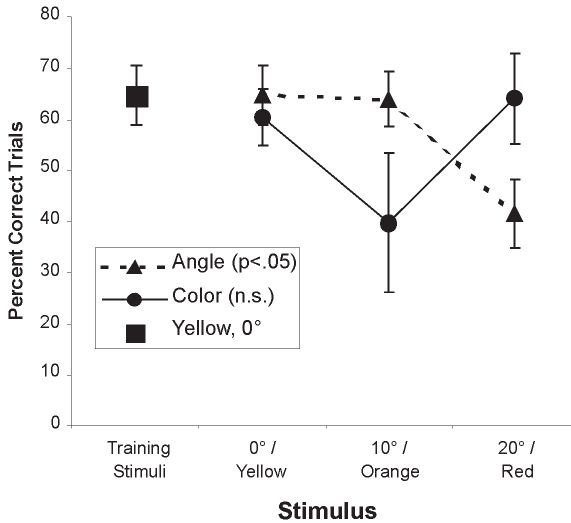


FIGURE 1 Percentage correct trials in Experiment 1 for training stimuli and two generalization dimensions: color (yellow, orange, and red) and angle (0° , 10° , and 20°). Note that generalization stimuli are averaged across the other dimension (e.g., 10° is the average of a red 10° , an orange 10° , and yellow 10°).

unrotated shapes. There was no effect of color, $Wald(1) < 1$. Finally, none of the two- or three-way interactions reached significance (all $p > .2$).

Given the absence of significant interactions, the data were collapsed along the color and orientation dimensions and two one-way ANOVAs were conducted for planned comparisons (see Figure 1). As expected, a one-way analysis of the color dimension showed no main effect, $F(2, 16) = 1.32, p > .2$,⁵ so no subsequent analyses were performed. A one-way analysis of orientation showed a significant main effect, $F(2, 16) = 9.33, p < .002$, and a decreasing linear trend as orientation moved away from 0° , $F(1, 8) = 12.52, p < .008$. Subsequent analyses revealed no difference between 0° and 10° stimuli ($F < 1$) and a difference between the average of these two and 20° , $F(1, 8) = 15.76, p = .004$.

Discussion

The results of Experiment 1 suggest that across the range of colors tested, infants' visual categories are not color specific: Squares and crosses can be identified correctly regardless of their color. This contrasts with Catherwood et al. (1989), who

⁵Note that a problem with the experimental control software created a design in which color and angle were partially confounded (there were no 0° red or orange stimuli). Given the significant effect of angle, the result of this confound should be an increased likelihood of detecting a false effect of color (due to angle). Despite this, there was no effect of color.

reported dominance for categorization by color over shape in 5-month-olds. Although color may be weighted more than shape when both are varied during training (see Experiment 3), given a training situation in which shape is the only dimension that is varied, color seems to play little role in generalization. These results also contrast with Colombo et al. (1990) by demonstrating that even when exposed to two shapes of a single color, infants can still categorize shapes despite variations in color. One interpretation of this difference is that MEH paradigms only allow the experimenter to specify irrelevant dimensions, whereas AEM allows the specification of the relevant dimensions.

In contrast to the results showing that color variations are ignored, shape categories appear to be orientation-specific—performance degrades with changes in orientation. Infants are, however, largely unaffected by a small amount of rotation (10°). This finding is consistent with other studies of orientation discrimination that have found boundaries for orientation categories to be between 7° and 15° of rotation from the vertical (Bornstein, 1982; Quinn & Bomba, 1986).⁶

EXPERIMENT 2: MULTIPLE-DIMENSIONAL CATEGORIES AND CUE WEIGHTING

Experiment 1 demonstrated that the variations in noncriterial dimensions can affect generalization responses to visual stimuli differing in shape. For 6-month-olds tested in the AEM paradigm, variations in color were not relevant to the shape-cued responses, but variations in orientation affected responding beyond a minimum threshold. Experiment 2 was designed to ask a related question: How are dimensions weighted when more than one dimension is criterial during training? To explore this question, infants were trained on stimuli that differed along two dimensions (orientation and spatial frequency) and then tested on stimuli that put these two dimensions into competition.

This particular generalization test may at first glance seem unnatural because, technically, there is no “right” answer. However, infants are often confronted with ambiguity arising from conflicting dimensional cues. In particular, with any multi-dimensional stimulus (typical of the real world), there are many potential dimensions that could be relevant to the definition of a category. Unless all dimensions are encoded (and weighted equally), it is quite reasonable to assume that there are

⁶However, in contrast to these discrimination experiments that used simple oriented bars or stimuli symmetrical in only one direction, it is not clear how to map orientation categories like vertical or horizontal onto the radially symmetric shapes used here (where there is no inherently preferred orientation). Thus, although previous research has established three basic orientation categories (vertical, horizontal, and oblique), and that the vertical and horizontal categories are the smallest (reviewed in Quinn & Eimas, 1987), the results reported here suggest that the vertical orientation category boundary corresponds to the minimum orientation displacement that affects shape categorization.

biases to generalize along some dimensions but not others. In particular, given a sparse conceptual space, there may not be a sufficient number of exemplars to unambiguously determine the relevant dimensions for categorization—inherent biases may be beneficial in reducing this space for more rapid learning.

Experiment 2 addressed this issue by using visual stimuli varying in both orientation and spatial frequency. A number of studies have established that infants categorize oriented bars into vertical, horizontal, and oblique categories (reviewed in Quinn & Eimas, 1987). None, however, has investigated how orientation categories might generalize across other dimensions. Moreover, given the well-established findings of orientation-specific spatial-frequency detectors in low-level visual processing by adults (de Valois & de Valois, 1988), determining the relative independence and weighting of spatial frequency and orientation in infants is an important goal.

We used two unique visual compounds for the training stimuli to determine the relative weighting of these dimensions in visual categorization. Two orientations and two spatial frequencies were selected: vertical and horizontal (oblique orientations were not used because of the reduced acuity for spatial frequency; see Leehy, Moskowitz-Cook, Brill, & Held, 1975), and square wave gratings with three or six black and yellow bars. The two training stimuli were three vertical bars (3V) and six horizontal bars (6H), and the generalization stimuli were the opposite combination of features: six vertical bars (6V) and three horizontal bars (3H).

Participants

Participants were 10 male and 9 female infants between 5 and 7 months old ($M = 26.1$ weeks, $SD = 2.92$). An additional 8 infants were excluded from the analysis due to fussiness or unreliable eye tracks. Recruitment was identical to Experiment 1.

Procedure

Participants were tested in the same manner as in Experiment 1 with several exceptions. Training consisted of 27 rather than 30 trials. Both stimuli consisted of a yellow circular patch containing a square wave grating of either three or six yellow bars that alternated with the black background. Both stimuli had a 9° maximum circular aperture yielding spatial frequencies of 3.6 and 1.7 cycles/degree. One stimulus contained three vertical bars (3V), and the other contained six horizontal bars (6H; see Figure 2). Stimulus- and reinforcement-side pairings were counterbalanced across participants (although for simplicity's sake we refer to 3V as being reinforced on the left and 6H on the right). These two training stimuli appeared in the same manner as in Experiment 1, and the same 16 reinforcers were used. The length of the delay period between the offset of the visual cue and the onset of the

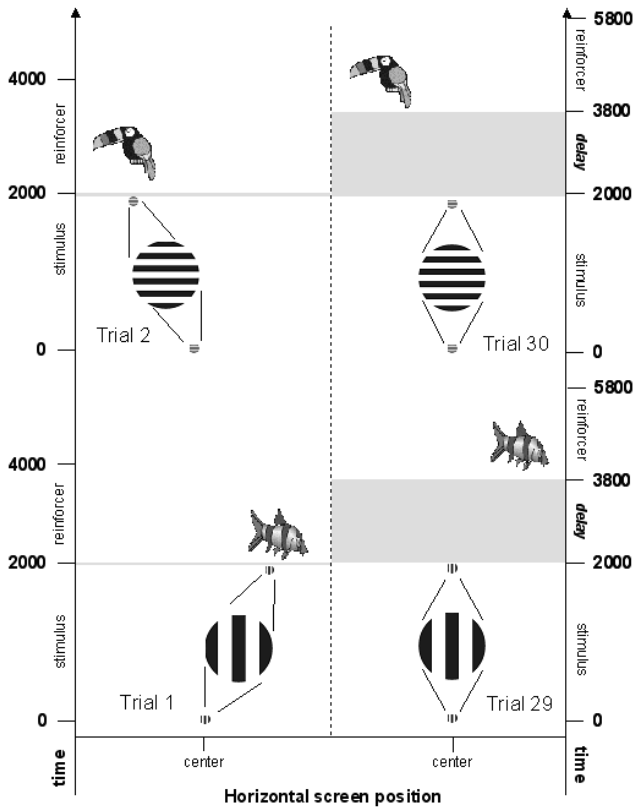


FIGURE 2 The display for Experiment 2. Time (within a trial) is represented vertically, and the horizontal position (on the screen) of stimulus items is shown on the horizontal axis. Early in the experiment (the left side, Trials 1 and 2), visual stimuli drift toward the location of their reinforcers, and there is no delay between the offset of the visual stimulus and the onset of the reinforcer. Gradually this drift decreases and the delay increases, until by the end of training (right panel, Trials 29 and 30) there is no drift and there is an 1,800-msec delay between stimulus offset and reinforcer onset.

visual reinforcer followed the same pattern as in Experiment 1 (increasing from 0–1,800 msec).

Unlike the previous experiment, the two visual cuing stimuli did not stay centered during training. As the stimulus loomed, it drifted a short distance in the direction toward which the reinforcer would appear (the maximum distance was 7°). As training progressed, this distance shortened until during testing, there was no directional cue to the reinforced side. Within the training block, the same animated displays used in Experiment 1 were played periodically (every four trials or so) to reengage the infant.

After 27 training trials, the infant and mother took a short break and then testing began. Odd-numbered testing trials presented the same 3V and 6H stimuli used during the training trials (with an 1,800-msec delay). Even numbered testing trials consisted of a novel stimulus (either 3H or 6V) followed by no reinforcer (although there was a subsequent delay equivalent to 1,800 msec plus the 2 sec during which the reinforcer would have been present). Testing proceeded until the infant was too fussy to continue, for an average of 18.9 trials. Data were scored in the same manner as Experiment 1.

Results

As in Experiment 2, each trial was scored as a look to the left if the infant spent more time looking to the left than to the right, and as a look to the right if he or she spent more time looking to the right during the 2,000 and 3,800 msec period after trial onset. A percentage correct could be assigned (for the training stimuli) to each participant based on these looks during testing. Overall, the infants performed quite well ($M = 72.3\%$, $SD = 0.23$), with 13 of 19 performing better than chance.

To explore each infant's generalization performance, we categorized infants as using one of four possible strategies during the testing phase (see Table 1). Of obvious interest were two strategies that reflected true generalization—categorizing by orientation or by the number of bars. Infants who used orientation as the dominant factor would be expected to look to the right for 3H (where 6H had been reinforced) and to the left for 6V (where 3V had been reinforced). Alternatively, infants who generalized along the spatial frequency dimension would look left for 3H (where 3V was reinforced) and right for 6V (where 6H was reinforced). The two remaining possible strategies—consistently looking to the left or right for all stimuli—would reflect a simple directional bias.

This scoring procedure enabled each trial to be counted for its contribution to the four strategies. For example, a trial in which the infant saw the 3H stimulus and

TABLE 1
The Four Generalization Strategies Explored in Experiment 2

<i>Stimulus</i>	<i>Orientation</i>	<i>Spatial Frequency</i>	<i>Left Biased</i>	<i>Right Biased</i>
3V	Left	Left	Training stimuli: Participants who were left or right biased did not learn and were excluded from analysis	
6H	Right	Right		
3H	Right	Left	Left	Right
6V	Left	Right	Left	Right

Note. 3V = three vertical bars; 6H = six horizontal bars; 3H = three horizontal bars; 6V = six vertical bars.

looked right contributed to both the orientation and the right-biased strategies. Because generalization by spatial frequency or orientation could be indicated by a look to either side on any given trial (depending on the stimulus), averaging across trials yields a picture of the dominant strategy for each infant.

Infants for whom the strategy with the most support was the orientation or spatial frequency strategy were considered generalizers. There were 7 such participants—the others appeared to adopt directionally biased strategies. Importantly, all of the infants who showed consistent generalization were part of the group that successfully learned to anticipate the appearance of the training stimuli. The independence of these two factors (learning and generalization) was assessed with a chi-square analysis, and they were found to be significantly nonindependent, $\chi^2(1, N = 19) = 5.12, p = .024$. This analysis underscores the important fact that in the anticipatory eye movement paradigm learning and generalization are linked—consistent generalization to novel stimuli reflects (and cannot occur without) learning of the original training stimuli.

We next explored whether individual infants generalized based on orientation or spatial frequency. We found a relatively equal distribution, with 3 infants preferring orientation and 4 preferring spatial frequency. This dimensional preference was not correlated with their performance on the training stimuli ($p > .2$) or with age ($p > .2$). These results suggest that by 6 months of age there are individual differences in the way spatial frequency and orientation contribute to the formation of visual categories.

Discussion

Experiment 2 yielded two important methodological results. First, the significant relation between learning and generalization validates the ability of the AEM paradigm to assess generalization after a brief training period with only two cuing stimuli. Second, individual infants appear to use different dimensions during generalization, suggesting that AEM has sufficient power to reveal individual differences. This absence of between-subjects consistency could be due to a developmental transition in which infants are moving from a preference for one dimension to a preference for another at 6 months of age. Alternatively, orientation and spatial frequency may never consistently dominate infants' encoding and categorization of multidimensional stimuli.

OCCLUSION-BASED AEM DISPLAYS

Although Experiments 1 and 2 demonstrated the utility of the AEM paradigm, several aspects could have been confusing or distracting to the infants. For ex-

ample, the delay between the disappearance of the cuing stimuli and the appearance of the reinforcer consisted of an unnatural block of “dead time” during which infants occasionally lost interest in the display (which was completely blank). Also, the gradually increasing delay may have partially extinguished anticipatory responses because of inhibition of return mechanisms (Clohessy, Posner, Rothbart, & Vecera, 1991).

Thus, Experiments 3 and 4 utilized a variant of the AEM paradigm that relies on a more natural configuration of moving objects and occlusion to motivate anticipatory eye movements. In this occlusion-based paradigm (see Figure 3 and <http://www.infancyarchives.com> for an example), each trial begins with a visual cuing stimulus that appears below an inverted-T-shaped occluder. The cuing stimulus moves behind the occluder and reemerges at either the top left or top right.

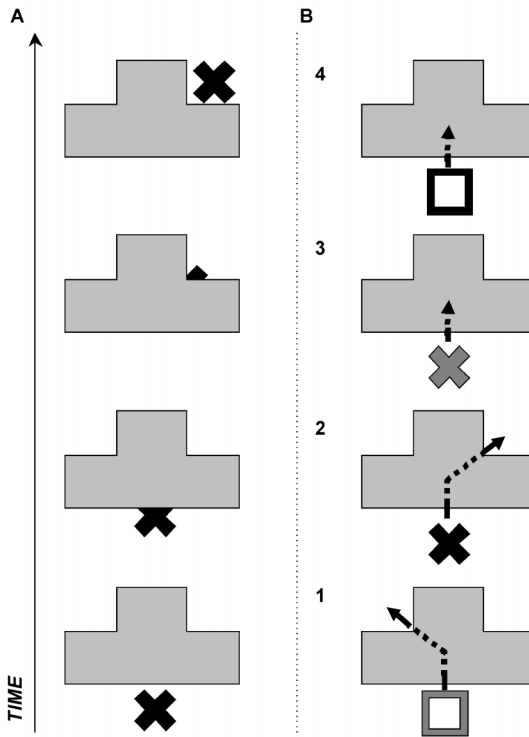


FIGURE 3 The occlusion-based AEM methodology used in Experiment 3. (A) The time course of a single trial: Visual stimuli start below the occluder, move behind it, and then emerge on either the right or left. (B) Stimulus used in Experiment 3: During training, red squares were reinforced (by reemerging) on the left, yellow crosses on the right (Items 1 and 2). During testing, these original stimuli were shown along with generalization stimuli (Items 3 and 4) in which a red cross or yellow square appeared and moved behind the occluder without reappearing.

The side of the occluder from which the cuing stimulus reappears is predicted by its identity: either visual identity (e.g., shape and color used in Experiment 3) or auditory identity (e.g., a sound that is presented concurrently with the appearance of a visual stimulus as in Experiment 4). After repeated exposure to two different cuing stimuli, infants learn to make anticipatory eye movements to the location where the visual stimulus will reappear.

This occlusion-based AEM paradigm offers several advantages over the disappearance paradigm used in Experiments 1 and 2. First, the display motivates anticipatory eye movements in a more natural, less arbitrary way. Rather than associating a cuing stimulus with a reward location that appears after a delay, this new display relies on the more common task of tracking a moving object behind an occluder. Previous research has shown that infants can anticipate the locations of regularly alternating visual stimuli in the visual expectation paradigm (Canfield et al., 1997; Haith et al., 1993) and can make anticipatory eye movements to the expected location of an object as it moves back and forth behind an occluder (Johnson, Amso, & Slemmer, 2003; see also Johnson, Slemmer, & Amso, 2004/*this issue*). In addition, the delay (during which anticipatory eye movements occur) is constant and consistent with the time the moving object would have spent behind the occluder if it were moving at a constant velocity. Thus both the training and testing phases are identical—the delay or position of the cuing stimuli does not change throughout the experiment, and the reappearance of the cuing stimulus from behind the occluder serves as the reinforcer.

EXPERIMENT 3: SHAPE AND COLOR CATEGORIZATION IN THE OCCLUSION PARADIGM

Experiment 3 was designed to be a relatively simple test of the occlusion-based AEM paradigm by comparing the relative weighting of shape and color in forming visual categories. Infants were trained on either red squares (which emerged on the left of the occluder) or yellow crosses (which emerged on the right). They were then tested on red crosses and yellow squares to determine whether shape or color would serve as the primary cue for category membership. Several previous studies have explored this question and have generally shown evidence for color primacy. As described earlier, Catherwood et al. (1989) demonstrated dishabituation for shapes with a different color than the habituation stimulus but not for a different shape with the same color. In addition, Colombo et al. (1990) showed a color preference when the shape did not vary during familiarization. This provides evidence that infants discriminate colors more easily than shapes and suggests that color may be a more heavily weighted dimension in categorization. However, later research by Catherwood (1994) demonstrated that with longer familiarization times infants could discriminate the shape categories, and Colombo et al. showed that

when color was varied during familiarization, infants could use shape for generalization. These results suggest that the bias for color may not entirely prevent the encoding and discrimination of shape.

Participants

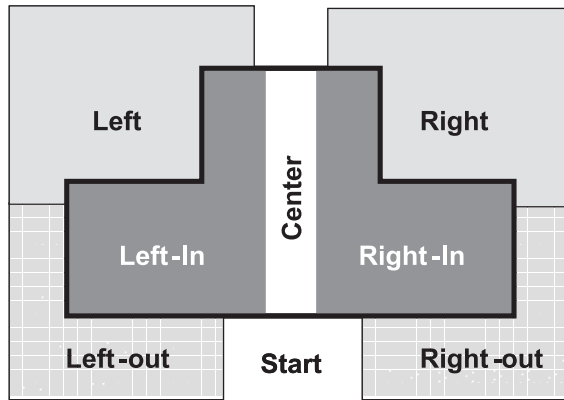
Participants consisted of 7 male and 15 female infants between 6 and 7 months of age ($M = 29.2$ weeks, $SD = 1.6$). Infants were recruited identically to the previous experiments.

Procedure

Participants were tested using the occlusion-based paradigm described earlier. Training was reduced from 27 to 16 randomly sequenced trials containing one of two cuing stimuli: a red square or a yellow cross. Both cuing stimuli expanded from a point at the bottom of the screen to 4° over a period of a second, shrank to half-size over the course of 500 msec, and then moved behind a light blue, inverted-T-shaped occluder.⁷ At a constant velocity behind the occluder, the cuing stimulus moved to the center point of the occluder and then diagonally to the top left or top right of the occluder where it reemerged (see Figure 3). Red squares moved to the top left, and yellow crosses moved to the top right. Both cuing stimuli were fully occluded for 750 msec. The onset of each trial was marked with a randomly selected sound (uncorrelated with the cuing stimuli) that helped to reorient the infant's gaze to the screen. Every four trials (during training and testing), the same animated displays from the earlier experiments were played periodically to reengage the infant.

Testing was conducted in blocks of four trials (interrupted by the reengaging stimulus). Each block of four trials contained (in random order) a red square and a yellow cross (the trained stimuli), and a yellow square and a red cross (the generalization stimuli). The pattern of stimulus motion for generalization stimuli was identical to the training stimuli except that the cuing stimulus did not emerge from behind the occluder (so as not to reinforce the infant) and the occluder remained on the screen during the time when the cuing stimulus would have reemerged. This random order of training and generalization trials (unlike the strict sequential or-

⁷Stimuli shrank to half-diameter before moving behind the occluder because the smaller size yielded more time behind the occluder (for a constant velocity). The inverted T shape of the occluder was designed to discourage infants from anticipating a straight trajectory (because the occluder "blocked" that location). This was only partially successful in that many infants made eye movements to this upper portion of the occluder during both training and testing. It is unclear, however, whether these eye movements were true anticipations or a general strategy of upward looking that reduced the offset between the change in gaze and the reappearance of the visual stimulus.



Off

FIGURE 4 Potential fixation locations used by coders for occlusion-based AEM experiments (3 and 4). Note that the center fixation location also includes fixations above the occluder but not clearly to the left or right of center.

dering in Experiments 1 and 2) was implemented to reduce perseveration responses in which (during a generalization trial) infants made anticipatory eye movements to the same side as the previous (reinforced) training trial. Infants were tested until they became too fussy to continue (an average of 17.6 trials).

Trained coders scored a videotaped record of the eye movements. Because of the greater spatial and temporal precision needed to accurately score these data, coders only relied on the output of the eye tracker. The larger video record was available to determine when and why track loss occurred and if the correct eye was being tracked. Fixations were categorized into one of nine different locations (see Figure 4), based on their rightward or leftward direction and also whether they were above, below, or inside the occluder. Fixations off the screen, obvious head turns away from the screen, and loss of the eye track were coded as “off.” Thus, there were nine coding areas to which infants could direct eye movements on each trial. The average correlation between coders for both temporal and spatial accuracy was 0.9 across both Experiments 3 and 4.

Results

Our first analyses assessed whether infants were able to learn the nonlinear trajectory of cue–stimulus motion. In spite of a long history of work demonstrating infants’ abilities to make anticipatory eye movements, von Hofsten, Feng, and Spelke (2000) reported that 6-month-olds have difficulty anticipating the motion

of objects that change trajectory by 90° under an occluder, even after repeated exposures. Thus, if infants were not capable of anticipating the side of emergence of the cuing stimulus from behind the occluder, we would expect very few fixations to the left or right stimulus locations, and the majority would be to the center location (which is on the upward linear path of motion). To test this, we computed the average number of eye movements per trial that fell within four categories that were collapsed from the nine coding locations: The right and left positions were classified as anticipatory eye movements, the right-in and left-in were also considered anticipations to the interior positions (as they would imply expectation of a nonlinear trajectory), the center position was the linear trajectory of the object's movement prior to occlusion, and the final category included eye movements off the screen or to the starting location of object movement.

A one-way ANOVA was conducted with the four collapsed categories of coding locations described earlier as a within-subjects factor. Three infants were excluded from this analysis because there were no eye movement data during the training trials in the testing phase (they were looking away), preventing us from assessing learning. The ANOVA revealed a marginal effect of coding category, $F(3, 54) = 2.413, p = .08$. A planned comparison between the center fixations (linear trajectory) and the anticipatory fixations (nonlinear trajectories) yielded no significant difference, $F(1, 17) = 1.468, p > .2$. Moreover, a one-sample t test indicated that the average number of anticipatory fixations (1.15 per trial) was greater than zero, $t(18) = 15.1, p < .001$. Thus, the linear trajectory "errors" reported by von Hofsten et al. (2000) did not dominate performance in this experiment.

The next analysis explored the relation between learning and generalization as in Experiment 1. First, each trial was scored as anticipating to the left (or right) if the infant spent more time looking to the left or left-in (or right or right-in) positions than to opposite positions. Left-out and right-out positions were not included in this analysis because they were not on the trajectory of the cuing stimulus as it moved from the starting position to either goal location. Because of the temporal component of the occlusion-based methodology, the critical time period was different for training trials than generalization trials. Training trials were scored between 2,000 msec and 3,650 msec. This was because the cuing stimulus became half-sized at 1,800 msec (at which point we could reasonably expect infants to start looking away from it) and oculomotor planning is known to take approximately 200 msec. At 3,450 msec the cuing stimulus was emerging from behind the occluder. The critical time period for generalization trials was longer (because the cuing stimulus never emerged), providing more time for anticipatory eye movements (from 2,000 msec to the end of the trial, 6,000 msec). A percentage correct was assigned (for the training stimuli) to each participant based on this trial-by-trial scoring. Overall, participants did not perform as well as in previous experiments ($M = 53.1\%$, $SD = 23\%$), with only 9 infants performing better than chance. This disappointing per-

formance is most likely due to the smaller number of training trials in this experiment (16 vs. 27 in Experiment 2).

To explore infants' generalization to the novel combinations of shape and color, we used the same strategic scoring technique as in Experiment 2. Each trial was characterized as consistent with a color strategy, a shape strategy, a left-biased strategy, or a right-biased strategy. There were 7 infants whose responses were consistent with a color or shape strategy (generalizers) and 14 infants who exhibited a bias strategy. Most important, 5 of the 7 generalizers also showed correct responding to the original training stimuli. Four infants did not provide either a generalization or a learning score (having looked away on those trials). A chi-square analysis of the remaining 18 infants did not reveal a significant relation between correct responding to the training stimuli and consistent (color or shape) generalization, $\chi^2(1, N = 23) = 2.10, p = .15$. However, after excluding 4 additional infants who had fewer than 10 test trials, the 6 remaining learners were generalizers, $\chi^2(1, N = 19) = 7.02, p = .008$.

When examining generalization preference among those infants who did not show a directional bias, a clear pattern of generalization occurred: All 5 babies who both generalized and learned showed a preference for color, as did 1 of the infants who generalized but did not learn. This number was higher than expected by chance, $\chi^2(1, N = 6) = 6.00, p = .014$. The remaining generalizer appeared to process the stimuli by shape but did not show evidence of learning and had fewer than 10 trials. Thus, overall, infants treated red crosses as if they came from the same category as red squares and yellow squares as if they came from the same category as yellow crosses.

Discussion

Experiment 3 demonstrated that infants are able to learn nonlinear trajectories for the movement of objects behind occluders. This contrasts with results from von Hofsten et al. (2000) and is consistent with those of Gredebäck and von Hofsten (2004/*this issue*). This more flexible pattern of tracking compared to the results from von Hofsten et al. is probably because the magnitude of their 90° shift in direction was twice as large as the 45° shift used in this experiment and larger than the circular trajectory used in Gredebäck and von Hofsten. In addition, similar to Gredebäck and von Hofsten, no instances of a linear trajectory ever occurred during the experiment. More important, Experiment 3 demonstrated that infants can also match the nonlinear trajectories behind an occluder with the visual characteristics of the cuing stimulus. The generalization data indicated that individual infants show consistent encoding of one of the two dimensions (color over shape), even when training creates a context in which either color or shape predicts the trajectory of the cuing stimulus. This replicates and extends earlier habituation results (Catherwood et al., 1989; Colombo et al., 1990, Experiment 3), and in turn vali-

dates the absence of a preference for color in Experiment 1 when shape was the only criterial dimension during training. Finally, these results suggest (contrary to Experiment 2) that tests of cue weighting can reveal significant between-subject dimensional biases in 6-month-olds.

EXPERIMENT 4: THE OCCLUSION-BASED PARADIGM, CATEGORY DEGRADATION, AND SPEECH

The final experiment examined categorization in the auditory modality to determine whether the occlusion-based AEM paradigm can be used to assess speech perception. Specifically, could infants identify words despite changes in duration and pitch? Infants were trained to look to the left after hearing *lamb* and to the right after hearing *teak*. These words were chosen because they were easily discriminable: *Lamb* contains a low vowel and continuant consonants, whereas *teak* contains a high vowel and plosive consonants (Slater, Quinn, Brown, & Hayes, 1999, used *mum* and *teat* in a cross-modal pairing task for the same reason). Pitch and duration were chosen because they are important features of infant-directed speech, which is typically slower and has higher pitch than adult-directed speech (Fernald et al., 1989). After training, infants were presented with novel tokens of these words that had artificially heightened pitch or lengthened duration to determine if they could overcome this variation to successfully identify the word.

Participants

Participants consisted of 12 male and 17 female infants between 5 and 7 months of age ($M = 24.4$ weeks, $SD = 1.8$). An additional 9 infants were not included in the analysis because they were too fussy to complete the session, and 1 was excluded because he was being raised bilingually. Participants were recruited in the same manner as in previous experiments.

Stimuli

Stimuli consisted of the words *lamb* and *teak* recorded by a female monolingual speaker of American English in an infant-directed speech register. Stimuli were recorded in a quiet room using a Kay Elemetrics 4300B CSL A/D system at a sample rate of 11,025 Hz. Both words were normalized to the same maximum amplitude to equate them for perceived loudness. Durations were 940 msec for *lamb* and 580 msec for *teak*. The average pitch for *lamb* was 218 Hz (range = 159–355 Hz). The average pitch for *teak* was 294 Hz (range = 175–356 Hz).

The modified speech tokens were constructed with the PSOLA analysis–resynthesis routines of the Praat speech analysis software. For duration modifications, the words were extended by 33% (to 1,207 msec for *lamb* and 756 msec for *teak*) and by 66% (to 1,489 msec for *lamb* and 915 msec for *teak*). For pitch modifications, resynthesis raised the original pitch by 20% and 40% (20%: 261 Hz for *lamb* and 334 Hz for *teak*; 40%: 305 Hz for *lamb* and 418 Hz for *teak*).

Procedure

Participants were tested in the same manner as in Experiment 3 with the following exceptions. Training consisted of 20 trials in which a visual cuing stimulus was paired with one of two words. The visual cuing stimulus was a circle of a color that was randomized across trials (to help maintain interest), rendering its visual properties irrelevant to the task of anticipating its reappearance on the right or left of the occluder. The colored circle expanded from a point at the bottom of the screen to 4° over a period of 1 sec. At this point of maximal size, an audio file of one of the two words was played: either the original (unaltered) *lamb* or *teak*. The colored circle then moved along one of the same two trajectories as in Experiment 3. The circle paired with *lamb* reemerged at the top left of the occluder, whereas the circle paired with *teak* reemerged at the top right. Every four trials (during training and testing), the same animated displays used in the earlier experiments were played periodically (every four trials or so) to reengage the infant.

Testing was conducted in blocks of four trials (punctuated by the reengaging stimulus). Each block of four trials contained (in random order) one unmodified *lamb* and *teak* stimulus and one (each) of the generalization *lamb* and *teak* stimuli (with either the heightened pitch or lengthened duration). This blocking of generalization trials maintained an equal number of trials in which a left and right anticipation was required (and rewarded) within each block of four so that short-term dependencies could not affect the data. As in Experiment 3, the visual stimulus did not emerge from behind the occluder for the generalization stimuli, and the delay was longer than during training trials. Infants were tested until they became too fussy to continue, for an average of 23.4 trials.

Trained coders scored the videotape by categorizing eye movements into the same nine display locations used in Experiment 3 (Figure 4). In this experiment, the output of the wide-angle video camera was replaced (on the tape) with the close-up image of the pupil (which was more useful in accounting for track loss; see Aslin & McMurray, 2004/*this issue*, Figure 2).

Results

To assess learning and generalization, each trial was scored as anticipating to the left (or right) if the infant spent more time looking to the left or left-in scoring loca-

tion (or to the right or right-in locations). As in Experiment 3, different critical time periods were used for training and generalization trials. Training trials were scored from word onset plus 200 msec of oculomotor planning (1,367 msec) to the reemergence of the colored circle from behind the occluder (3,400 msec). Generalization trials were scored from the same onset to the end of the trial (5,400 msec).

Overall, participants' performance on the training stimuli in Experiment 4 was similar to Experiment 3 ($M = 52.4\%$, $SD = 21\%$), with 11 of 29 infants performing better than chance. To explore the effect of pitch and duration on infants' generalization, we used the same logistic regression technique as in Experiment 1 to examine generalization trials from the 11 infants who performed better than chance. Two separate logistic regression analyses (one for pitch and one for duration) included effects for word (*lamb* or *teak*) and three steps of pitch or duration (the original unmodified stimuli, plus the two modified ones). In both cases the dependent variable was binary and indicated whether the infant looked to the "correct" side on that trial.

We first looked at the effect of pitch and word (*lamb* or *teak*; see Figure 5). There was a significant effect of word, $Wald(1) = 5.93$, $p = .015$, with infants performing better on *lamb* than *teak*. There was no effect of pitch, $Wald(1) = 0.086$, $p > .2$, and no interaction with word, $Wald(1) = 1.92$, $p > .15$. The analysis by duration yielded no effect of word, $Wald(1) = 0.96$, $p > .2$, or interaction of word with duration, $Wald(1) = 1.69$, $p > .15$, suggesting that the differences between words in the previous analysis were carried by the generalization stimuli. There was, however, a highly significant (negative) effect of duration, $Wald(1) = 9.54$, $p = .002$: As

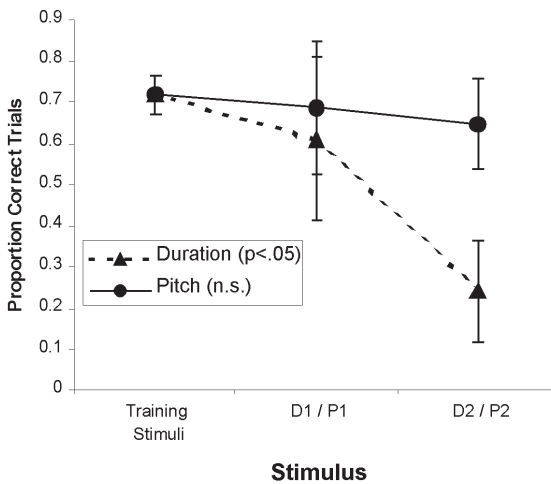


FIGURE 5 Categorization performance as a function of duration (D1, D2) and pitch (P1, P2) in Experiment 4.

the duration of the stimuli increased, performance declined. Post-hoc tests revealed that this effect was due to a difference between the unmodified stimulus and the second step: unmodified versus 66% longer, $Wald(1) = 8.04, p = .005$; unmodified versus 33%, $Wald(1) = 0.89, p > .2$.

Paired *t* tests verified the results found using logistic regression. Because of a number of missing trials, the two pitch-modified and the two duration-modified stimuli were averaged to create a single score for pitch-modified stimuli and a single score for duration-modified stimuli. Both were compared to the unmodified stimuli. There was no significant effect of pitch, $t(9) = 0.61, p > .2$, but there was a significant effect of duration, $t(8) = 3.60, p = .007$, confirming the results of the logistic regression.

Discussion

Experiment 4 demonstrated that infants can map arbitrary auditory inputs onto visual trajectories and that this can be used to determine how speech (and other) sounds are categorized. Although the majority of the infants did not learn this mapping, more than one third showed clear evidence of learning and reliable generalization. Patterns of generalization indicated that infants were successful at maintaining word identification despite pitch variation, but their performance was negatively affected by variation in duration. These results suggest that global temporal cues (e.g., duration or amplitude envelope) may be more important than spectral cues in infants' word recognition.

GENERAL DISCUSSION

In our series of four experiments, we demonstrated the utility of the AEM paradigm for assessing the encoding and categorization of a variety of visual and auditory stimuli. AEM can reveal which sources of information infants use to encode the inherent multidimensionality of visual and auditory stimuli and which dimensions are weighted most heavily in judging the similarity of novel exemplars to the two training stimuli. The AEM paradigm also has the potential to be used for studies of individual differences in how infants attend to and weight dimensional information.

The occlusion-based AEM paradigm used in Experiments 3 and 4 also demonstrated an important fact about the infant's expectations for trajectories: Given enough exposure, infants can learn nonlinear trajectories under occlusion (in contrast to von Hofsten et al., 2000). However, the occlusion-based technique also showed a decrement in learning: Whereas most babies were able to learn the training stimuli in Experiments 1 and 2, Experiments 3 and 4 showed much less success. This suggests that the 16 training trials used in Experiment 3 and

the 20 in Experiment 4 may not have been enough—25 or 30 may be more realistic. Although the rate of attrition using this small, fixed number of training trials was high, it does not differ from other conditioning paradigms (e.g., HAS or conditioned head turning).

Despite this rate of attrition, we believe that the occlusion-based AEM paradigm will prove even more useful in the future. It not only offers a much simpler training regimen, as there is no need to vary the duration of the anticipation period or location of the stimuli, but it has the potential when coupled with online eye tracking to tailor the training regimen for each infant. That is, the experimenter could make a determination online (during the experiment) to terminate (or extend) training to the two cuing stimuli before moving into the generalization phase. This more flexible training algorithm might further reduce participant attrition and capture more repeated measurements for infants who have learned more rapidly than expected.

The occlusion-based AEM paradigm also seems to capture the infants' interest more effectively than the disappearance paradigm used in Experiments 1 and 2. A comparison across all experiments of the amount of "off-display" looking time yielded a significant difference between the older AEM methodologies in Experiments 1 and 2 ($M = 191$ video frames/trial) and the occlusion-based ones in Experiments 3 and 4 ($M = 143$ video frames/trial), $t(78) = 4.5, p < .001$. Conversely, occlusion-based methodologies yielded more looks to target areas than the older methodology ($M_{\text{occlusion}} = 136$ video frames; $M_{\text{older}} = 83$), $t(76) = 7.23, p < .001$. Finally, due to the relatively natural visual experience provided by the occlusion-based method, this paradigm can also be used (unaltered) to test adults and older children who may find the "visual reinforcers" to be contrived. Work in our lab has already used this technique successfully to examine speech categorization in adults (McMurray, Maye, Lathrop, & Aslin, 2003) and yielded similar results to more traditional phoneme decision methods. Thus, this same methodology may be unique in that it can provide developmental information across a broad age range.

Although one of the strengths of the AEM paradigm is that it taps the natural variation within categories (as opposed to "teaching" infants what variation to include in a category), some of the results here suggest that the training procedure can influence the resulting categories. One could argue that by training on only a single token within each category, infants may simply be computing similarity to that exemplar rather than to the category per se. Although this should be addressed empirically, our replication of the results obtained with MEH by Colombo et al. (1990) suggests that this is not the case. Moreover, exemplar models of adult categorization (e.g., Nosofsky & Johansen, 2000) argue that this computation of similarity to exemplars may be fundamental to categorization. AEM, then, may provide an extremely useful tool for exploring the biases and abilities that infants bring to bear on these processes. Thus, future work should vary the types of training materials (perhaps using multiple exemplars for each category)

to explore the nature of this dynamic acquisition of categories in the lab versus categorization strategies or dimensional biases that infants may bring to the lab based on prior experience.

In conclusion, AEM has been shown to be a successful addition to existing methodologies for the study of infant categorization. Unlike other methodologies, it provides a behavioral label for each exemplar, allows for the collection of many repeated measures per infant, and does not require training on the structure of the category before testing. In many ways, AEM is an infant analogue to the 2AFC task that is ubiquitous in adult research and represents an important methodological tool for the study of infant categorization.

ACKNOWLEDGMENTS

Bob McMurray is now at the Department of Psychology, University of Iowa.

This research was supported by National Institutes of Health Research Grant HD-37082 to Richard N. Aslin. We thank Ellen Graubard for contributions in the design and data collection of Experiment 2, Elizabeth Gramzow, Koleen Gochal-McCrink, and Julie Markant for exceptional assistance in organizing the collection and coding of data, and the students in the Infant Perception Lab for coding the videotapes.

REFERENCES

- Aslin, R. N., & McMurray, B. (2004). Automated corneal-reflection eye tracking in infancy: Methodological developments and applications to cognition. *Infancy*, *6*, 155-163.
- Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception* (pp. 67-96). New York: Academic.
- Bomba, P. (1984). The development of orientation categories between 2 and 4 months of age. *Journal of Experimental Child Psychology*, *37*, 609-636.
- Bornstein, M. H. (1982). Perceptual anisotropies in infancy: Ontogenetic origins of inequalities in spatial vision. In H. W. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior* (pp. 77-123). New York: Academic.
- Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). The categories of hue in infancy. *Science*, *191*, 201-202.
- Canfield, R. L., Smith, E. G., Brezsnayak, M. P., & Snow, K. L. (1997). Information processing through the first year of life. *Monographs of the Society for Research in Child Development*, *62*(Serial No. 250).
- Catherwood, D. (1994). Exploring the seminal phase in infant memory for color and shape. *Infant Behavior and Development*, *17*, 235-243.
- Catherwood, D., Crassini, B., & Freiberg, K. (1989). Infant response to stimuli of similar hue and dissimilar shape: Tracing the origins of the categorization of objects by hue. *Child Development*, *60*, 752-762.

- Clohessy, A. B., Posner, M. I., Rothbart, M. K., & Vecera, S. P. (1991). The development of inhibition of return in early infancy. *Journal of Cognitive Neuroscience*, 3, 345–350.
- Cohen, L. (Ed.). (2000). Perceptually based approaches to understanding early categorization [Thematic collection]. *Infancy*, 1, 29–122.
- Colombo, J., McCollam, K., Coldren, J. T., Mitchell, D. W., & Rash, S. J. (1990). Form categorization in 10-month-olds. *Journal of Experimental Child Psychology*, 49, 173–188.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9, 133–150.
- de Valois, R. L., & de Valois, K. K. (1988). *Spatial vision*. Oxford, England: Oxford University Press.
- Fagan, J. F. (1976). Infants' recognition of invariant features of faces. *Child Development*, 47, 627–638.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mother's and father's speech to preverbal infants. *Journal of Child Language*, 16, 477–501.
- Fodor, J. A., Garrett, M. F., & Brill, S. (1975). Pi ka pu: The perception of speech sounds by prelinguistic infants. *Perception & Psychophysics*, 18, 74–78.
- Gredebäck, G., & von Hofsten, C. (2004). Infants' evolving representations of object motion during occlusion: A longitudinal study of 6- to 12-month-old infants. *Infancy*, 6, 165–184.
- Haith, M. M., Wentworth, N., & Canfield, R. L. (1993). The formation of expectations in early infancy. In C. Rovee-Collier & L. Lipsitt (Eds.) *Advances in infancy research* (Vol. 8, pp. 251–297). Norwood, NJ: Ablex.
- Hayne, H. (1996). Categorization in infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 10, pp. 79–120). Norwood, NJ: Ablex.
- Husain, J., & Cohen, L. (1981). Infant learning of ill-defined categories. *Merrill-Palmer Quarterly*, 27, 443–456.
- Johnson, S., Amso, D., & Slemmer, J. (2003). Development of object concepts in infancy: Evidence for early learning in an eye tracking paradigm. *Proceedings of the National Academy of Sciences, USA*, 100, 10568–10573.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, 6, 185–201.
- Jusczyk, P., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, 23, 648–654.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, 66, 1668–1679.
- Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, 6, 263–285.
- Leehy, S. C., Moskowitz-Cook, A., Brill, S., & Held, R. (1975). Orientational anisotropy in infant vision. *Science*, 190, 900–902.
- Mandler, J. M. (1998). Representation. In W. Damon (Series Ed.) and D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (pp. 255–308). New York: Wiley.
- McMurray, B., Maye, J., Lathrop, A., & Aslin, R. N. (2003, November). *Categorical perception of speech: Task variations in infants and adults*. Paper presented at the Auditory Perception, Cognition, and Action Meeting, Vancouver, British Columbia, Canada.
- McMurray, B., Spivey, M., & Aslin, R. (2000). The perception of consonants by adults and infants: Categorical or categorized? *The Working Papers in the Language Sciences at the University of Rochester*, 2, 215–256 [Electronic journal].
- Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of “multiple system” phenomena in perceptual categorization. *Psychonomic Bulletin and Review*, 7, 375–402.
- Quinn, P., & Bomba, P. (1986). Evidence for a general category of oblique orientations in four-month-old infants. *Journal of Experimental Child Psychology*, 42, 345–354.

- Quinn, P., & Eimas, P. D. (1987). On categorization in early infancy. In J. Oates & S. Sheldon (Eds.), *Cognitive development in infancy* (pp. 131–161). Hove, England: Lawrence Erlbaum Associates, Ltd.
- Quinn, P., Slater, A., Brown, E., & Hayes, R. (2001). Developmental change in form categorization in early infancy. *British Journal of Developmental Psychology, 19*, 207–218.
- Slater, A., Quinn, P., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory–visual pairing. *Developmental Science, 2*, 333–338.
- von Hofsten, C., Feng, Q., & Spelke, E. (2000). Object representation and predictive action in infancy. *Developmental Science, 3*, 193–205.