

# Enhanced Attention to Speaking Faces Versus Other Event Types Emerges Gradually Across Infancy

Lorraine E. Bahrick and James Torrence Todd  
Florida International University

Irina Castellanos  
The Ohio State University

Barbara M. Sorondo  
Florida International University

The development of attention to dynamic faces versus objects providing synchronous audiovisual versus silent visual stimulation was assessed in a large sample of infants. Maintaining attention to the faces and voices of people speaking is critical for perceptual, cognitive, social, and language development. However, no studies have systematically assessed when, if, or how attention to speaking faces emerges and changes across infancy. Two measures of attention maintenance, habituation time (HT) and look-away rate (LAR), were derived from cross-sectional data of 2- to 8-month-old infants ( $N = 801$ ). Results indicated that attention to audiovisual faces and voices was maintained across age, whereas attention to each of the other event types (audiovisual objects, silent dynamic faces, silent dynamic objects) declined across age. This reveals a gradually emerging advantage in attention maintenance (longer HTs, lower LARs) for audiovisual speaking faces compared with the other 3 event types. At 2 months, infants showed no attentional advantage for faces (with greater attention to audiovisual than to visual events); at 3 months, they attended more to dynamic faces than objects (in the presence or absence of voices), and by 4 to 5 and 6 to 8 months, significantly greater attention emerged to temporally coordinated faces and voices of people speaking compared with all other event types. Our results indicate that selective attention to coordinated faces and voices over other event types emerges gradually across infancy, likely as a function of experience with multimodal, redundant stimulation from person and object events.

*Keywords:* infant attention development, faces and voices, faces versus objects, audiovisual and visual stimulation, selective attention

*Supplemental materials:* <http://dx.doi.org/10.1037/dev0000157.supp>

The nature and focus of infant attention and its change across age is a critically important topic, as attention provides a foundation for subsequent perceptual, cognitive, social, and language development. Selective attention to information from objects and events in the environment provides the basis for what is perceived, and what is perceived provides the basis for what is learned and, in turn, what is remembered (Bahrick & Lickliter, 2012, 2014; Gibson, 1969). What we attend to shapes neural architecture and dictates the input for learning and, in turn, further perceptual and cognitive development (Bahrick & Lickliter, 2014; Greenough &

Black, 1992; Knudsen, 2004). Faces and voices and audiovisual speech constitute one class of events thought to be highly salient to young infants and preferred over other event types (Doheny, Hurwitz, Insoft, Ringer, & Lahav, 2012; Fernald, 1985; Johnson, Dziurawiec, Ellis, & Morton, 1991; Sai, 2005; Valenza, Simion, Cassia, & Umiltà, 1996; Walker-Andrews, 1997). Moreover, attention to the rich, dynamic, and multimodal stimulation in face-to-face interaction is fundamental for fostering cognitive, social, and language development (Bahrick, 2010; Bahrick & Lickliter, 2014; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001; Mundy &

Lorraine E. Bahrick and James Torrence Todd, Department of Psychology, Florida International University; Irina Castellanos, Department of Otolaryngology – Head and Neck Surgery, The Ohio State University; Barbara M. Sorondo, Florida International University Libraries, Florida International University.

This research was supported by grants from the National Institute of Child Health and Human Development (R01 HD053776, R03 HD052602, and K02 HD064943), the National Institutes of Mental Health (R01 MH062226), and the National Science Foundation (SLC SBE0350201) awarded to Lorraine E. Bahrick. Irina Castellanos and Barbara M. Sorondo were supported by National Institutes of Health/National Institute of General Medical Sciences Grant R25 GM061347. A portion of these data was presented at the 2009 and 2011 biennial meetings of the Society for

Research in Child Development, the 2009 and 2010 annual meetings of the International Society for Autism Research, and the 2008 annual meeting of the International Society for Developmental Psychobiology. We gratefully acknowledge Melissa Argumosa, Laura C. Batista-Taran, Ana C. Bravo, Yael Ben, Ross Flom, Claudia Grandez, Lisa C. Newell, Walueska Pallais, Raquel Rivas, Christoph Ronacher, Mariana Vaillant-Molina, and Mariana Werhahn for their assistance in data collection, and Kasey C. Soska and John Colombo for their constructive comments on the manuscript.

Correspondence concerning this article should be addressed to Lorraine E. Bahrick, Department of Psychology, Florida International University, 11200 SW 8 Street, Miami, FL 33199. E-mail: [bahrick@fiu.edu](mailto:bahrick@fiu.edu)

Burnette, 2005; Rochat, 1999). However, we know little about how or when attention to dynamic audiovisual faces and voices emerges and develops, or about the salience of speaking faces and voices relative to other event types, and how this relative salience changes across infancy. The present study assessed the emergence of infant attention across 2 to 8 months of age to dynamic audiovisual faces and voices of people speaking compared with silent visual faces, audiovisual object events, and silent object events in a large sample of 801 infants.

Infant attention to social events, particularly faces and voices of caretakers, scaffolds typical social, cognitive, and language development. The rich naturalistic stimulation from faces and voices provides coordinated multimodal information not available in unimodal visual stimulation from faces, or auditory stimulation from voices alone (Bahrick, 2010; Bahrick & Lickliter, 2012, 2014; Gogate & Hollich, 2010; Mundy & Burnette, 2005). Social interaction relies on rapid coordination of gaze, voice, and gesture, and infants detect social contingencies in multimodal dyadic synchrony (Feldman, 2007; Harrist & Waugh, 2002; Jaffe et al., 2001; Rochat, 1999, 2007; Stern, 1985). Contingent responses to infant babbling promote and shape speech development (Goldstein, King, & West, 2003). Mapping words to objects entails coordinating visual and auditory information, and parents scaffold learning by timing verbal labels with gaze and/or object movement (Gogate, Bahrick, & Watson, 2000; Gogate & Hollich, 2010; Gogate, Maganti, & Bahrick, 2015; Gogate, Walker-Andrews, & Bahrick, 2001). Parents also exaggerate visual and auditory prosody to highlight meaning-bearing parts of the speech stream ("multimodal motherese"; Gogate et al., 2000, 2015; Kim & Johnson, 2014; Smith & Strader, 2014). These activities require careful attention and differentiation of signals in the face and voice of the caretaker. However, most research on the emergence of attention to faces and voices has focused on attention to static faces, silent dynamic faces, and on voices devoid of faces, whereas little research has systematically assessed the emergence of attention to naturalistic coordinated faces and voices of people speaking across infancy, or compared attention to faces of people speaking with that of objects producing naturalistic sounds.

The multimodal stimulation provided by the synchronous faces and voices of people speaking also provides intersensory redundancy. Intersensory redundancy is highly salient to infants and attracts attention to properties of stimulation that are redundantly specified (i.e., amodal properties). For example, amodal information, such as tempo, rhythm, duration, and intensity changes, is available concurrently and in temporal synchrony across faces and voices during speech. Sensitivity to these properties provides a cornerstone for early social, cognitive, and language development, underlying the development of basic skills such as detecting word-referent relations, discriminating native from non-native speech, and perceiving affective information, communicative intent, and social contingencies (see Bahrick & Lickliter, 2000, 2012; Gogate & Hollich, 2010; Lewkowicz, 2000; Walker-Andrews, 1997; Watson, Robbins, & Best, 2014). Detection of redundancy provided by amodal information is considered the "glue" that binds stimulation across the senses (Bahrick & Lickliter, 2002, 2012). The intersensory redundancy hypothesis (Bahrick, 2010; Bahrick & Lickliter, 2000, 2012, 2014), a theory of selective attention, describes how the salience of intersensory redundancy bootstraps early social development by attracting and maintaining infant attention to

coordinated stimulation (e.g., faces, voices, gesture, and audiovisual speech) from unified multimodal events (as opposed to unrelated streams of auditory and visual stimulation), a critical foundation for typical development.

Relative to nonsocial events, social events provide an extraordinary amount of intersensory redundancy. "Social" stimuli are typically conceptualized as involving people or animate objects; however, definitions have varied. Some researchers have included static images of faces and face-like patterns, dolls, or nonhuman animals as social stimuli (see Farroni et al., 2005; Ferrara & Hill, 1980; Legerstee, Pomerleau, Malcuit, & Feider, 1987; Maurer & Barrera, 1981; Simion, Cassia, Turati, & Valenza, 2001). Here, "social events" are defined as "people events" including dynamic faces and voices of people speaking or performing actions, whereas "nonsocial events" are defined as "object events" in which people are not visible or readily apparent. Social events are typically more variable, complex, and unpredictable than nonsocial events, and subtle changes conveying meaningful information occur rapidly (Adolphs, 2001, 2009; Bahrick, 2010; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Jaffe et al., 2001). For example, communicative exchanges involve interpersonal contingency and highly intercoordinated and rapidly changing temporal, spatial, and intensity patterns across face, voice, and gesture (Bahrick, 2010; Gogate et al., 2001; Harrist & Waugh, 2002; Mundy & Burnette, 2005). Intersensory redundancy available in temporally coordinated faces and voices guides infant attention to, and promotes early detection of, affect (Flom & Bahrick, 2007; Walker-Andrews, 1997), word-object relations in speech (Gogate & Bahrick, 1998; Gogate & Hollich, 2010), and prosody of speech (Castellanos & Bahrick, 2007). Moreover, faces and voices are processed more deeply (e.g., event related potential [ERP] evidence of greater reduction in amplitude of late positive slow wave) and receive more attentional salience (e.g., greater amplitude of Nc component) when they are synchronized compared with when they are asynchronous or visual alone (Reynolds, Bahrick, Lickliter, & Guy, 2014).

Prior studies investigating the development of attention, mostly to static images or silent events, have found, in general, that attention becomes more flexible and efficient across infancy with decreases in look length and processing time and concurrent increases in the number of looks away from stimuli (Colombo, 2001; Johnson, Posner, & Rothbart, 1991; Ruff & Rothbart, 1996). Courage, Reynolds, and Richards (2006) investigated attention to silent dynamic faces across multiple ages. They found greater attention (longer looks, greater heart rate change) to dynamic events than to static images and to social events (Sesame Street scenes and faces) than to achromatic patterns across 3 to 12 months of age. Although attention to all event types declined from 3 through 6 months, attention to social (but not nonsocial) events increased from 6 through 12 months of age, illustrating the salience of silent, dynamic faces. In contrast, in a longitudinal study of 1.5- to 6.5-month-olds, Hunnius and Geuze (2004) found that the youngest infants showed a greater percentage of looking time to silent scrambled than unscrambled faces, and older infants showed longer median fixations to silent scrambled than unscrambled faces. Thus, the development of attention to silent speaking faces versus nonface events remains unclear.

Only a few studies have investigated the development of attention to multimodal social events. Reynolds, Zhang, and Guy

(2013) found greater attention (average look duration) to audiovisual events (both synchronous and asynchronous) than visual events at 3 and 6 months (but not 9 months), and greater attention to Sesame Street scenes than to geometric black-and-white patterns at all ages. Further, attention to Sesame Street decreased across age, whereas attention to geometric patterns remained low and constant. Findings suggest that audiovisual, complex scenes are more salient to infants at certain ages than simple patterns and unimodal events. However, Shaddy and Colombo (2004) found that although look duration decreased from 4 to 6 months of age, infants showed only marginally greater attention to dynamic faces speaking with sound than without sound. These studies indicate that, overall, looking time declines across 3 to 6 months (as attention becomes more efficient), and that infant attention may be best maintained by dynamic, audiovisual events. However, there is no consensus regarding the extent to which, or conditions under which, infants prefer to attend to speaking faces over object events and over silent events, or whether preferences are apparent in early development or emerge gradually across infancy. A systematic investigation across age is needed.

To begin to address these important questions, we conducted a large-scale systematic study of the emergence and change in infant attention to both audiovisual and visual speaking faces and moving objects. Our primary focus was to chart the early emergence and change across 2 to 8 months in attention maintenance to synchronous faces and voices compared with other event types. We included 2-month-olds in order to capture the early emergence of attentional patterns for speaking faces. Only one study (Hunnus & Geuze, 2004) had assessed infants as young as 2 months, and several others had found attentional differences to faces versus other events by 3 months of age. We also focused on the relative interest in audiovisual face events compared with each of the other event types at each age, and describe its change across age, given that the distribution of attention to different events provides the input for perceptual, cognitive, and social development.

Attention maintenance was assessed according to two complementary measures (typically assessed separately): habituation time (HT) and look-away rate (LAR). HT indexes overall looking time prior to reaching the habituation criterion and is one of the most commonly used measures in infant attention and perception. HT is thought to reflect the amount of time it takes to process or encode a stimulus (Bornstein & Colombo, 2012; Colombo & Mitchell, 2009; Kavšek, 2013). Processing speed improves across age, and faster processing predicts better perceptual and cognitive skills and better cognitive and language outcomes (Bornstein & Colombo, 2012; Colombo, 2004; Rose, Feldman, Jankowski, & Van Rossem, 2005). LAR reflects the number of looks away from the habituation stimulus per minute. In contrast with HT, developments in infant LAR and more general attention shifting behaviors (e.g., anticipatory eye movements, visual orienting) are thought to reflect early self-regulation and control of attention (Colombo & Cheatham, 2006; Posner, Rothbart, Sheese, & Voelker, 2012), and are predictive of inhibitory control (Pérez-Edgar et al., 2010; Sheese, Rothbart, Posner, White, & Fraundorf, 2008) and the regulation of face-to-face interactions (Abelkop & Frick, 2003). LAR has also been used as a measure of distractibility or sustained attention to a stimulus (Oakes, Kannass, & Shaddy, 2002; Pérez-Edgar et al., 2010; Richards & Turner, 2001). Regardless of the underlying processes, together, measures of HT and LAR provide

independent yet complementary indices of attention maintenance—the attention-holding power of a stimulus. The attention-holding power of faces and voices of people speaking is of high ecological significance, given their importance for scaffolding cognitive, social, and language development.

We assessed these indices of infant attention to dynamic audiovisual and visual silent speaking faces and moving objects in a sample of 801 infants from 2 to 8 months of age by recoding data from habituation studies collected in our lab. We expected that an attention advantage for faces over objects would emerge gradually across development as a function of infants' experience interacting with social events and the salience of redundant audiovisual stimulation compared to nonredundant visual stimulation. We were thus interested in whether and at what age infants would show, first, greater attention maintenance (longer HT and lower LAR) to faces than objects, consistent with an emerging "social preference," and second, greater attention to audiovisual than visual stimulation, consistent with findings of the attentional salience of intersensory redundancy. Third, we predicted that infants would show increasingly greater attention maintenance to audiovisual speaking faces than to each of the other event types (audiovisual objects, visual faces, and visual objects) across age. Fourth, consistent with prior studies, we expected these effects to be evident in the context of increased efficiency of attention across age, with overall declines in HT and increases in LAR across age.

## Method

### Participants

Eight hundred one infants (384 females and 417 males) between 2 and 8 months of age participated in one of a variety of studies conducted between the years 1998 and 2009 (see Table 1). Each infant participated in only a single habituation session and in no other concurrent studies. Infants were categorized into four age groups: 2-month-olds ( $n = 177$ ; 86 females;  $M = 70.58$  days,  $SD = 8.40$ ), 3-month-olds ( $n = 210$ ; 106 females;  $M = 97.03$ ,  $SD = 12.92$ ), 4- to 5-month-olds ( $n = 247$ ; 112 females;  $M = 145.32$ ,  $SD = 13.05$ ), and 6- to 8-month-olds ( $n = 167$ ; 80 females;  $M = 210.34$ ,  $SD = 28.62$ ). Seventy-eight percent of the infants were Hispanic, 13% were Caucasian of non-Hispanic origin, 3% were African American, 1% were Asian, and 5% were of unknown or mixed ethnicity/race. All infants were healthy and born full-term, weighing at least 5 pounds, and had an Apgar score of at least 9.

### Stimulus Events

The stimulus events consisted of videotaped displays of dynamic faces of people speaking and objects impacting a surface presented under conditions of audiovisual (with their natural, synchronized soundtracks) or visual (same events with no soundtrack) stimulation. The events had been created for a variety of studies, both published and unpublished (see Table 1), and were chosen to provide consistency across age and condition (visual, audiovisual) in the event types included. Videos of speaking faces depicted the head and upper torso of an unfamiliar woman or man, with gaze directed toward the camera, reciting a nursery rhyme or a story with slightly or very positive affect, using infant-directed speech

Table 1  
*Composition of the Data Set (N = 801) Broken Down as a Function of the Source of the Data (Published Studies, Conference Presentations, and Unpublished Data)*

	<i>n</i>
Published studies	
Bahrck, Lickliter, Castellanos, and Todd (2015)	53
Bahrck, Lickliter, and Castellanos (2013)	80
Bahrck, Lickliter, Castellanos, and Vaillant-Molina (2010)	48
Bahrck, Lickliter, and Flom (2006)	69
Bahrck and Lickliter (2004)	68
Bahrck, Flom, and Lickliter (2002)	16
Bahrck and Lickliter (2000)	37
Total	371
Conference presentations	
Newell, Castellanos, Grossman, and Bahrck (2009, March)	49
Bahrck, Shuman, and Castellanos (2008, March)	35
Bahrck, Newell, Shuman, and Ben (2007, March)	48
Bahrck et al. (2005, April)	32
Bahrck et al. (2005, November)	16
Castellanos, Shuman, and Bahrck (2004, May)	50
Bahrck, Lickliter, Shuman, Batista, and Grandez (2003, April)	19
Total	249
Unpublished data	181
Grand total	801

*Note.* Procedures involved the presentation of a control event (toy turtle), a minimum of six and maximum of 20 infant-control habituation trials (defined by a 1-s or 1.5-s look-away criterion), a maximum trial length of 45 s or 60 s, and a habituation criterion defined as a 50% reduction in looking on two successive trials relative to the infant's own looking level on the first two trials (baseline) of habituation. See reference list for complete references.

(see Bahrck, Lickliter, & Castellanos, 2013). Fifty-five percent of the faces were of Hispanic individuals, 35% were Caucasian, and 10% were of unknown race/ethnicity. Videos of moving objects primarily depicted a red toy hammer striking a surface in one of several distinctive rhythms and tempos, producing naturalistic impact sounds (see Bahrck, Flom, & Lickliter, 2002; Bahrck & Lickliter, 2000), sometimes accompanied by a synchronous light flashing or yellow baton tapping (Bahrck, Lickliter, Castellanos, & Todd, 2015). Videos for a secondary/comparison data set (see Footnote 2) depicted metal or wooden single or multiple objects suspended from a string, striking a surface in an erratic temporal pattern (see Bahrck, 2001). However, these events were not included in the main event set because they were not represented in the oldest age group.

## Apparatus

The stimulus events were played using Panasonic video decks (AGDS545 and AGDS555 or AG6300 and AG7550) and were displayed on a color TV monitor (Sony KV-20520 or Panasonic BT-S1900N). Participants sat approximately 55 cm from the screen in a standard infant seat. All soundtracks were presented from a centrally located speaker. Two experimenters (a primary and secondary observer), occluded by a black curtain behind the TV screen, measured infants' visual fixations by pressing buttons on a button box or a gamepad connected to a computer that collected the data online. Data from the primary observers were used for analyses. Data from the secondary observers were used to

calculate interobserver reliability. Average Pearson's product moment correlations between the judgments of the primary and secondary observers were .98 ( $SD = .02$ ; range = .95 to .99).

## Procedure

**Selection of data sets.** The data were compiled from a variety of infant-controlled habituation studies (see Habituation procedure) conducted between 1998 and 2009 (see Table 1). Data sets were selected so that stimuli would be relatively consistent across age and within event type (speaking faces, moving objects). The audiovisual and visual stimuli were also quite similar, given that most data sets were designed to compare the same stimuli in audiovisual versus visual conditions. We excluded data sets with stimuli depicting people manipulating objects. These criteria allowed for a relatively clean comparison of attention across conditions (see Table 2).

**Habituation procedure.** All infants participated in a variant of the infant-control habituation procedure in which they were habituated to a single video depicting an audiovisual or a visual face speaking or object moving. Habituation trials commenced when the infant looked at the screen and lasted until the infant looked away for 1 s or 1.5 s, or until the maximum trial length of 45 s or 60 s had elapsed (for details, see Bahrck et al., 2013; Bahrck, Lickliter, Castellanos, & Vaillant-Molina, 2010). Infants first viewed a control event (video of a toy turtle whose arms spun, making a whirring sound), followed by a minimum of six and a maximum of 20 habituation trials. The habituation criterion was defined as a 50% decrease in looking time during two consecutive trials relative to the infant's own mean looking time across the first two (i.e., baseline) habituation trials. Following habituation, there were two no-change posthabituation trials, a series of test trials, and a final presentation of the control event (see Bahrck, 1992, 1994, for further details). Data were used from only the habituation portion (and not the test or control trials) for participants who passed the habituation and fatigue (looking on final control trial greater than 20% of looking on initial control trial) criteria. The raw data records were rescored to derive two primary measures of attention (a third measure was also scored and is summarized in Footnote 3).

**Indices of attention.** Two measures of attention maintenance were calculated for each infant and then averaged across infants: HT and LAR. HT was calculated as the total number of seconds an infant spent looking across all habituation trials. LAR per minute was calculated as the total number of times an infant looked away

Table 2  
*Number of Participants as a Function of Age in Months, Event Type (Faces, Objects), and Type of Stimulation (Audiovisual, Visual)*

Age in months	Faces		Objects		Total
	Audiovisual	Visual	Audiovisual	Visual	
2	52	36	34	55	177
3	98	44	33	35	210
4-5	52	41	94	60	247
6-8	74	25	34	34	167
Total	276	146	195	184	801

(defined as 0.2 s or greater) from the stimulus event during habituation, divided by HT, and multiplied by 60.

HT and LAR were examined for outliers of three standard deviations or greater with respect to each cell mean (Age  $\times$  Event Type  $\times$  Type of Stimulation). There were 11 outliers for HT (1% of the sample) and 11 for LAR (1% of the sample). Given that these scores were likely to bias estimates of cell means, and that they constituted a small percentage of the sample, they were removed from subsequent analyses.

## Results

Results for HT and LAR are presented in Table 3. These measures (together and individually) are conceptualized as an index of attention maintenance or attention holding value of the stimulus events. Greater attention maintenance is reflected by longer HT and lower LAR.

### Attention to Faces and Objects as a Function of Age (2, 3, 4–5, 6–8 Months) and the Type of Stimulation (Audiovisual, Visual)

To examine the development of attention to face and object events under conditions of redundant audiovisual and nonredundant visual stimulation as a function of age, ANOVAs with age (2, 3, 4–5, 6–8 months), event type (faces, objects), and type of stimulation (bimodal audiovisual, unimodal visual) as between-subjects factors were conducted for HT and LAR. After conducting overall analyses, we followed up with analyses at each age to determine at what age any simple effects and interactions were evident. Further, planned a priori comparisons were conducted to explore the nature of interactions, and all used a modified, multi-stage Bonferroni procedure to control the familywise error rate for multiple comparisons (Holm, 1979; Jaccard & Guilamo-Ramos, 2002).

Results of ANOVAs demonstrated significant main effects of age, event type, and type of stimulation for both measures.<sup>1</sup> Consistent with predictions and prior research indicating increased efficiency and flexibility in attention across infancy, main effects of age indicated shorter HT,  $F(3, 774) = 22.06, p < .001, \eta_p^2 = .08$ , and higher LAR,  $F(3, 752) = 46.15, p < .001, \eta_p^2 = .16$ , to events overall with increasing age (see Figures 1b and 2b, respectively). Further, main effects of event type indicated significantly longer HT,  $F(1, 774) = 40.07, p < .001, \eta_p^2 = .05$ , and lower LAR,  $F(1, 752) = 6.01, p = .01, \eta_p^2 = .01$ , for faces than objects (see Figures 1c and 2c), indicating an overall social preference.<sup>2</sup> Finally, a main effect of type of stimulation revealed significantly longer HT,  $F(1, 774) = 28.15, p < .001, \eta_p^2 = .04$ , and lower LAR,  $F(1, 752) = 59.70, p < .001, \eta_p^2 = .07$ , to audiovisual than visual stimulation (see Figures 1d and 2d), consistent with the proposed attentional salience of intersensory redundancy. There was no significant three-way interaction between age, event type, and type of stimulation ( $ps > .34$ ). However, the main effects were each qualified by important interactions (see Face Versus Object Events, Audiovisual Versus Visual Events, and Audiovisual Face Events) and thus should be considered in the context of these interactions.<sup>3</sup>

### Face Versus Object Events: Attention to Faces (Compared With Objects) Increases Across Age

For both measures, significant interactions between age and event type indicated longer HT,  $F(3, 774) = 3.54, p = .01, \eta_p^2 = .01$ , and lower LAR,  $F(3, 752) = 6.42, p < .001, \eta_p^2 = .03$ , for faces than objects emerged across age (see Figure 1c and 2c). Planned comparisons revealed no difference in either HT or LAR for faces versus objects at 2 months, but longer HT and lower LAR for faces than objects was evident at older ages ( $ps < .02$ , Cohen's  $d$  range = .42 to 1.02; except no difference in LAR at 4 to 5 months). Differences in HT between faces and objects showed a dramatic increase across age, with a mean difference of 7.27 s (2%) at 2 months, to a mean difference of 74.61 s (30%) at 6 to 8 months of age (see Figure 1c). Findings demonstrate increasing differences in attention maintenance to faces across development, with 2-month-olds demonstrating no significant difference in attention to faces versus objects, and 6- to 8-month-olds showing the greatest attentional advantage for faces (however, this advantage is carried by attention to audiovisual faces, see Audiovisual Face Events).

### Audiovisual Versus Visual Events: Attention to Audiovisual (Compared With Visual) Stimulation Increases With Age

For both measures, significant interactions of age and type of stimulation were apparent, revealing increasingly longer HT,  $F(3, 774) = 3.84, p = .01, \eta_p^2 = .02$ , and lower LAR,  $F(3, 752) = 10.61, p < .001, \eta_p^2 = .04$ , to audiovisual than visual events across age (see Figures 1d and 2d). Planned comparisons revealed longer HT to audiovisual than visual events at 2 months ( $p = .02, d = .34$ ; but no difference for LAR), no differences for either measure at 3 months ( $ps > .50$ ; instead, infants showed greater attention to faces than objects), and longer HT and lower LAR to audiovisual than visual stimulation at 4 to 5 and 6 to 8 months ( $ps < .001, d$  range = .64 to 1.00). Similar to attention to face versus object

<sup>1</sup> Additional analyses were performed to assess the roles of participant gender (female, male) and ethnicity (Hispanic, not Hispanic). Results indicated no significant main effects of ethnicity on HT or LAR ( $ps > .39$ ), and no main effects of gender on LAR ( $p = .28$ ). A significant main effect of gender on HT emerged,  $F(1, 758) = 11.89, p = .001$ , with longer HT for males ( $M = 182.70, SD = 108.20$ ) than females ( $M = 156.20, SD = 107.94$ ). However, because there were no significant interactions of gender or ethnicity with other factors (age, event type, or type of stimulation;  $ps > .26$ ), we chose not to include gender or ethnicity in subsequent analyses.

<sup>2</sup> To explore generalization to a broader class of nonsocial events, we analyzed a secondary data set ( $N = 175$ ) depicting different nonsocial stimuli (single and compound objects impacting a surface in an erratic temporal pattern; see Bahrick, 2001). These events had primarily been presented in audiovisual conditions, and were not presented to infants in the oldest age category (6–8 months), and thus they did not meet criteria for inclusion in our main data set. However, when these data were merged with those of the main data set, results of ANOVAs indicated no change in significance levels for main effects or interactions for any of the variables.

<sup>3</sup> A third variable, average length of look (ALL), was calculated by dividing HT by the number of looks away. Analyses indicated the results of ALL mirrored those of HT with decreasing ALL across age, longer ALL to faces than objects by 3 months of age, and longer ALL to audiovisual faces by 6 to 8 months of age. Further, ALL was highly correlated with HT,  $r = .47, p < .001$ . For additional details, see the online supplemental materials.

**Table 3**  
*Means and Standard Deviations (in Parentheses) for Two Indices of Attention as a Function of Age, Event Type (Faces, Objects), and Type of Stimulation (Audiovisual, Visual) for the Entire Habituation Phase*

Index/age	Audiovisual faces	Visual faces	Audiovisual objects	Visual objects	Faces	Objects	Audiovisual	Visual	Overall
<b>Habituation time</b>									
2 months	225.20 (95.46)	177.15 (94.78)	205.65 (110.63)	182.15 (114.86)	201.18 (95.12)	193.90 (112.75)	215.43 (103.05)	179.65 (104.82)	197.54 (103.93)
3 months	245.18 (123.24)	228.24 (113.47)	163.49 (69.45)	179.94 (116.17)	236.71 (118.36)	171.72 (92.81)	204.34 (96.35)	204.09 (114.82)	204.21 (105.58)
4-5 months	219.64 (142.90)	134.19 (71.89)	142.68 (67.30)	115.90 (57.55)	176.92 (107.40)	129.29 (62.43)	181.16 (105.10)	125.05 (64.72)	153.10 (84.91)
6-8 months	223.78 (128.98)	101.91 (57.10)	98.20 (46.17)	78.27 (41.95)	162.85 (93.04)	88.24 (44.06)	160.99 (87.58)	90.09 (49.53)	125.54 (68.55)
Mean across age	228.45 (122.65)	160.37 (84.31)	152.51 (73.39)	139.07 (82.63)	194.41 (103.48)	145.79 (78.01)	190.48 (98.02)	149.72 (83.47)	170.10 (90.74)
<b>Look-away rate</b>									
2 months	6.22 (3.50)	7.49 (4.15)	5.77 (2.71)	6.21 (3.15)	6.86 (3.83)	5.99 (2.93)	6.00 (3.11)	6.85 (3.65)	6.42 (3.38)
3 months	6.61 (3.62)	6.91 (3.85)	8.03 (3.13)	8.65 (4.27)	6.76 (3.74)	8.34 (3.70)	7.32 (3.38)	7.78 (4.06)	7.55 (3.72)
4-5 months	8.43 (4.93)	12.76 (6.75)	8.14 (4.16)	12.29 (4.74)	10.60 (5.84)	10.22 (4.45)	8.29 (4.55)	12.53 (5.75)	10.41 (5.15)
6-8 months	7.43 (5.04)	12.52 (6.36)	10.76 (4.13)	15.23 (5.89)	9.98 (5.70)	13.00 (5.01)	9.10 (4.59)	13.88 (6.13)	11.49 (5.36)
Mean across age	7.17 (4.27)	9.92 (5.28)	8.18 (3.53)	10.60 (4.51)	8.55 (4.78)	9.39 (4.02)	7.67 (3.90)	10.26 (4.90)	8.97 (4.40)

*Note.* Habituation time is the total number of seconds to habituation. Look-away rate is the number of looks away per minute.

events, differences in HT to audiovisual versus visual stimulation became increasingly apparent across age, with a mean difference of 35.78 s (9%) at 2 months, and a mean difference of 70.90 s (28%) at 6 to 8 months of age (see Figure 1d). These findings demonstrate that by 2 months, and becoming more apparent across age, infants show enhanced attention maintenance to audiovisual events providing naturalistic synchronous sounds compared with visual events.

**Audiovisual Face Events: Differences in Attention to Audiovisual Faces (Compared With the Other Three Types of Stimulation) Increase With Age**

Finally, consistent with our predictions, a significant interaction emerged between event type and type of stimulation for HT,  $F(1, 774) = 12.65, p < .001, \eta_p^2 = .01$  (but not LAR,  $p = .63$ ; see Table 3). Planned comparisons revealed that collapsed across age, infants showed greater HT and lower LAR to audiovisual faces than to audiovisual objects, visual faces, and visual objects ( $ps < .03, d$  range = .26 to .85). These effects were then explored at each age to address our main prediction.

To characterize the emergence of attention to faces versus objects as a function of type of stimulation, tests of simple effects and interactions were conducted at each age (2, 3, 4-5, 6-8 months) with event type (faces, objects) and type of stimulation (audiovisual, visual) as between-subjects factors, followed by planned comparisons (controlling for family wise error) to determine the nature of any event type and type of stimulation interactions at each age. Means and standard deviations are presented in Table 3 (see also Figures 1a and 2a).

At 2 months of age, main effects of event type for both HT and LAR failed to reach significance ( $ps = .64$  and  $.21$ , respectively), indicating no difference in attention maintenance to faces versus objects. In contrast, a significant main effect of type of stimulation for HT,  $F(1, 774) = 5.36, p = .02, \eta_p^2 = .01$  (but not LAR,  $p = .21$ ), indicated greater attention maintenance to audiovisual than visual events. No planned comparisons were significant at 2 months. At 3 months, significant main effects of event type emerged for both measures, indicating longer HT,  $F(1, 774) = 17.92, p < .001, \eta_p^2 = .02$ , and lower LAR,  $F(1, 752) = 5.42, p = .02, \eta_p^2 = .01$ , to faces than objects, with no differences in attention as a function of type of stimulation ( $ps = .99$  and  $.50$ , respectively). Thus, by 3 months, infant attention is best maintained by dynamic faces (over objects) regardless of type of stimulation. Planned comparisons for HT revealed greater attention to audiovisual faces than audiovisual objects and visual objects ( $ps < .001$  and  $= .001$ , respectively,  $d$  range =  $.54$  to  $.82$ ; but not visual faces,  $p = .36$ ), suggesting the above effect was carried by audiovisual faces.

Enhanced attention maintenance to audiovisual faces over each of the other three event types emerged at 4 to 5 months and was most evident at 6 to 8 months. At 4 to 5 months, a significant interaction between event type and type of stimulation was evident for HT,  $F(1, 774) = 4.78, p = .03, \eta_p^2 = .00$ , with longer HT to audiovisual faces than to each of the other event types ( $ps < .001, d$  range =  $.69$  to  $.95$ ). LAR results indicated only a main effect of type of stimulation,  $F(1, 752) = 49.29, p < .001, \eta_p^2 = .06$ , with greater attention to audiovisual faces than visual faces and objects ( $ps < .001, d$  range =  $.73$  to  $.80$ ), but not audiovisual

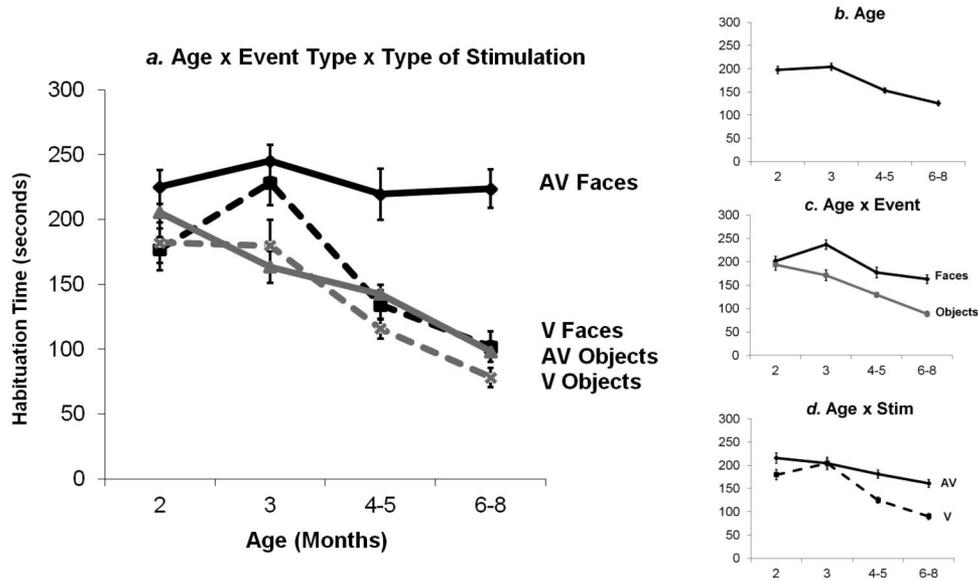


Figure 1. Mean habituation times (HTs) as a function of (a) age, event type (faces, objects), and type of stimulation (audiovisual, visual); (b) age; (c) age and event type; and (d) age and type of stimulation. Figure 1a depicts HT to audiovisual faces (AV Faces), audiovisual objects (AV Objects), visual faces (V Faces), and visual objects (V Objects). Error bars depict standard error of the mean.

objects ( $p = .72$ ). At 6 to 8 months, the Event  $\times$  Stimulation interaction for HT was still significant,  $F(1, 774) = 8.97, p = .003, \eta_p^2 = .01$ . Planned comparisons again revealed longer HT ( $ps < .001, d$  range = .65 to .85) to audiovisual faces than all other event types (39% greater than for audiovisual objects,

37% for visual faces, and 48% for visual objects). Planned comparisons also revealed lower LAR ( $ps < .001, d$  range = .72 to 1.42) to audiovisual faces than to all other event types by 6 to 8 months (18% lower than for audiovisual objects, 25% for visual faces, and 34% for visual objects).

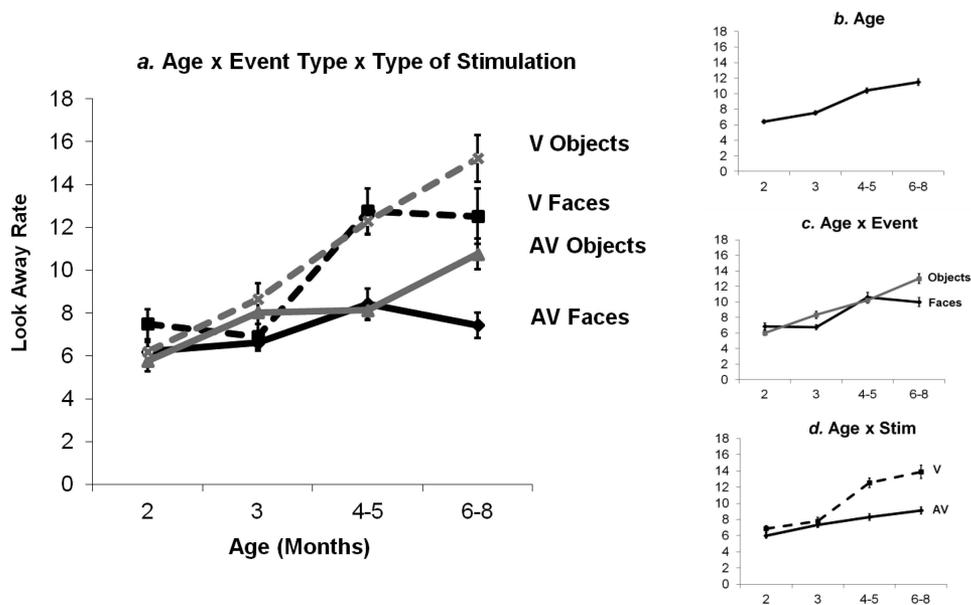


Figure 2. Mean look-away rate (LAR) as a function of (a) age, event type (faces, objects), and type of stimulation (audiovisual, visual); (b) age; (c) age and event type; and (d) age and type of stimulation. Figure 2a depicts LAR to audiovisual faces (AV Faces), audiovisual objects (AV Objects), visual faces (V Faces), and visual objects (V Objects). Error bars depict standard error of the mean.

Thus, consistent with our predictions, attention was best maintained by speaking faces that were both audible and visible. At 3 months, this trend was emerging for HT with greater attention to audiovisual faces than two of the three other event types (visual faces and objects). However, heightened attention to audiovisual faces over each of the other three event types was clearly evident by 4 to 5 months for HT, and was most evident at 6 to 8 months of age for both HT and LAR. These findings demonstrate clear evidence of an attentional advantage, across both measures, to audible and visible face-voice events emerging across age.

### Developmental Trajectories: Characterizing the Nature of Change in Attention to Face and Object Events Across Age

Linear regression analyses were conducted to reveal the slopes of attention across age (increase, decrease, or no change) for each of the event types and to more specifically address the nature of the attentional advantage, which emerged across age, for audiovisual faces over each of the other event types (audiovisual objects, visual faces, visual objects). Was the attentional advantage a result of declining attention to each of the other three event types and constant or increasing attention to audiovisual faces across age? If so, we would expect significant differences between slopes of attention across age for audiovisual faces and all other event types. Linear regression analyses of HT ( $R^2 = .19$ ),  $F(7, 782) = 26.43$ ,  $p < .001$ , and LAR ( $R^2 = .22$ ),  $F(7, 760) = 30.29$ ,  $p < .001$ , with age as a continuous variable revealed little to no change in attention to audiovisual faces across age, but significant linear changes in attention for each of the three other conditions (audiovisual objects, visual faces, visual objects; see Table 4, Figure 3).<sup>4</sup> Specifically, for audiovisual faces, there was no change across age in attention for HT ( $p = .67$ ). In

contrast, slopes for HT to each of the three other event types showed a sharp linear decrease across age ( $ps < .001$ ). Moreover, the slope for HT to audiovisual faces was significantly different from the slopes for each of the other event types (audiovisual objects, visual faces, visual objects;  $ps < .01$ ), with no differences at 2 months and dramatic differences by 6 to 8 months. Thus, the attentional advantage for audiovisual faces over each of the other event types emerges across 2 to 8 months of age as a result of decreasing looking time to audiovisual objects, visual faces, and visual objects across age, but a maintenance of high levels of looking to audiovisual faces across age.

Similarly, although the LAR to audiovisual faces showed only a slight, marginally significant increase across age (a .34 average increase per month;  $p = .08$ ), the slopes for LAR for all other event types showed a sharp linear increase across age ( $ps < .001$ ). The slope for audiovisual faces for LAR was significantly different from that of visual faces and visual objects ( $ps < .01$ ), but unlike that of HT, it was not different from that of audiovisual objects ( $p = .19$ ). In fact, the slope for LAR for audiovisual objects was also significantly different from that of visual faces and visual objects ( $ps < .001$ ). Thus, for LAR, slopes for visual events (both objects and faces) increased sharply with age, whereas slopes for audiovisual events showed significantly less change across age. It was only by the age of 6 months that LAR for audiovisual faces was significantly different from that for audiovisual object events ( $p = .03$ ). This pattern suggests that the development of attention as indexed by LAR parallels that of HT, but differences in attention to audiovisual faces from each of the other event types emerges slightly later than for HT. Thus, the attentional advantage for audiovisual faces for LAR emerges across age as a result of significant increases in the rates of looking away from audiovisual objects, visual faces, and visual objects across age, but a low level of looking away from audiovisual faces, with only a slight, marginal increase across 2 to 8 months of age.

These analyses illustrate that although attention to audiovisual faces was maintained across age, attention to the other three event types decreased systematically across age. This results in an increasing disparity across age in selective attention to audiovisual faces compared with each of the other event types. Given that attention to audiovisual, speaking faces serves as a foundation for cognitive, social, and language development (Bahrick & Lickliter, 2014), the relative distribution of attention at a given age to audiovisual faces compared with other event types is of central importance. This distribution reflects the product of selective attention and forms the input and foundation for later development. To capture this emphasis, we depict the data as proportions (proportion of attention allocated to each event type with respect to overall attention across event types at each age; see Figures 4a and 4b). Proportions (HT and LAR) for each participant within each event type at each age were calculated with respect to the total HT and LAR across the four event types at each age. This reflects the relative distribution of attention maintenance to each of the four

Table 4

Raw Scores: Results From Regression Analysis Assessing Slopes Across Age for the Two Measures of Attention (Habituation Time, Look-Away Rate) as a Function of Event Type (Faces, Objects) and Type of Stimulation (Audiovisual, Visual)

Measure	<i>b</i> estimate	SE	<i>p</i> value	<i>b</i> *
<b>Habituation time</b>				
Overall	-16.49	2.19	<.001	-.26
Faces	-7.18	3.55	.04	-.11
Objects	-17.78	2.68	<.001	-.28
Audiovisual	-13.90	2.88	<.001	-.22
Visual	-22.22	3.19	<.001	-.35
Audiovisual faces	-2.18	4.19	.60	-.03
Audiovisual objects	-16.39	3.80	<.001	-.26
Visual faces	-27.31	6.19	<.001	-.43
Visual objects	-19.70	3.60	<.001	-.31
<b>Look-away rate</b>				
Overall	1.05	.12	<.001	.31
Faces	.63	.16	<.001	.22
Objects	1.09	.13	<.001	.38
Audiovisual	.56	.12	<.001	.19
Visual	1.60	.14	<.001	.56
Audiovisual faces	.34	.19	.08	.12
Audiovisual objects	.67	.17	<.001	.23
Visual faces	1.63	.28	<.001	.57
Visual objects	1.61	.17	<.001	.56

Note. *b* estimate = unstandardized regression coefficient; SE = standard error of the unstandardized coefficient; *b*\* = standardized regression coefficient.

<sup>4</sup> We also assessed whether slopes across age would be better characterized by a quadratic or cubic function. Analyses indicated only a significant quadratic function for one variable, visual objects for LAR ( $p = .04$ ), with a steeper increase at younger than older ages. However, the difference between linear versus quadratic models was virtually zero ( $\Delta R^2 < .01$ ), indicating no significant gain by using a quadratic model.

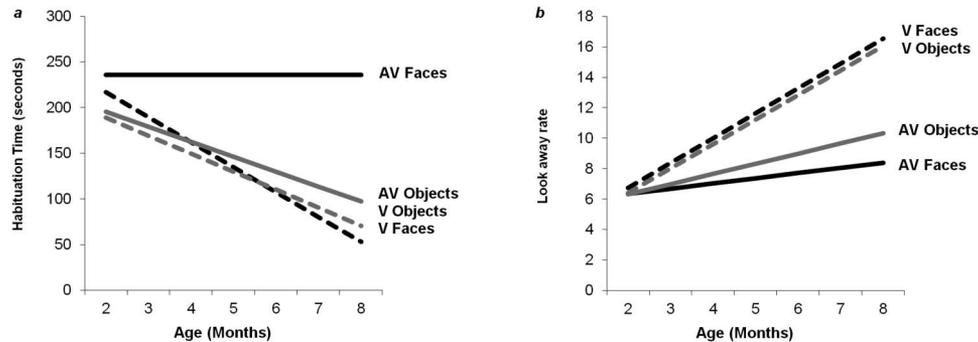


Figure 3. Best-fitting regression lines depicting change across age in attention maintenance to four event types (audiovisual faces, visual faces, audiovisual objects, visual objects) for (a) habituation time (HT), and (b) look-away rate (LAR). Figures 3a and 3b depict HT and LAR, respectively, to audiovisual faces (AV Faces), audiovisual objects (AV Objects), visual faces (V Faces), and visual objects (V Objects).

event types at each age. As is evident from the proportion scores, the proportion of attention maintenance to audiovisual faces relative to other event types *increases systematically* across age. Regression analyses of HT ( $R^2 = .20$ ,  $F(7, 782) = 27.00$ ,  $p < .001$ , and LAR ( $R^2 = .10$ ,  $F(7, 760) = 11.95$ ,  $p < .001$ , proportion scores with age as a continuous variable (see Figures 4c and 4d, and Table 5) revealed a sharp linear increase in HT ( $p < .001$ ) and a sharp linear decrease in LAR ( $p < .001$ ), reflecting increased distribution of attention maintenance to audiovisual faces across age with respect to total attention across event types at each age.<sup>5</sup> In contrast, the proportion of attention to visual objects decreased systematically across age ( $ps < .01$ ), and there was no change in attention maintenance to audiovisual objects or visual faces ( $ps > .17$ ). Moreover, slopes for the proportions of HT and LAR to audiovisual faces were significantly different from those of other event types (audiovisual objects, visual faces, visual object;  $ps < .03$ ).<sup>6</sup>

Together, the two sets of regression analyses indicate (a) that attention maintenance to audiovisual speaking faces remains high and constant across 2 to 8 months of age, whereas maintenance to all other event types decreases with age; and (b) that the proportion of time infants selectively attend to audiovisual faces compared with the other event types at each age increases across 2 to 8 months of age.

### Correlations Between HT and LAR

Finally, we conducted correlational analyses between HT and LAR both overall and as a function of age. Collapsed across age, HT and LAR were significantly, negatively correlated ( $r = -.51$ ,  $p < .001$ ), with shorter HT associated with higher LAR. Further, HT-LAR correlations increased with age, from  $r = -.24$  ( $p < .01$ ) at 2 months to  $r = -.62$  ( $p < .001$ ) at 6 to 8 months, suggesting that overall looking time and looking away become more tightly coupled with age. Thus, relations between the two indices of attention grew stronger with age, suggesting greater consistency with shorter looking time and more frequent looking away across age.

### Discussion

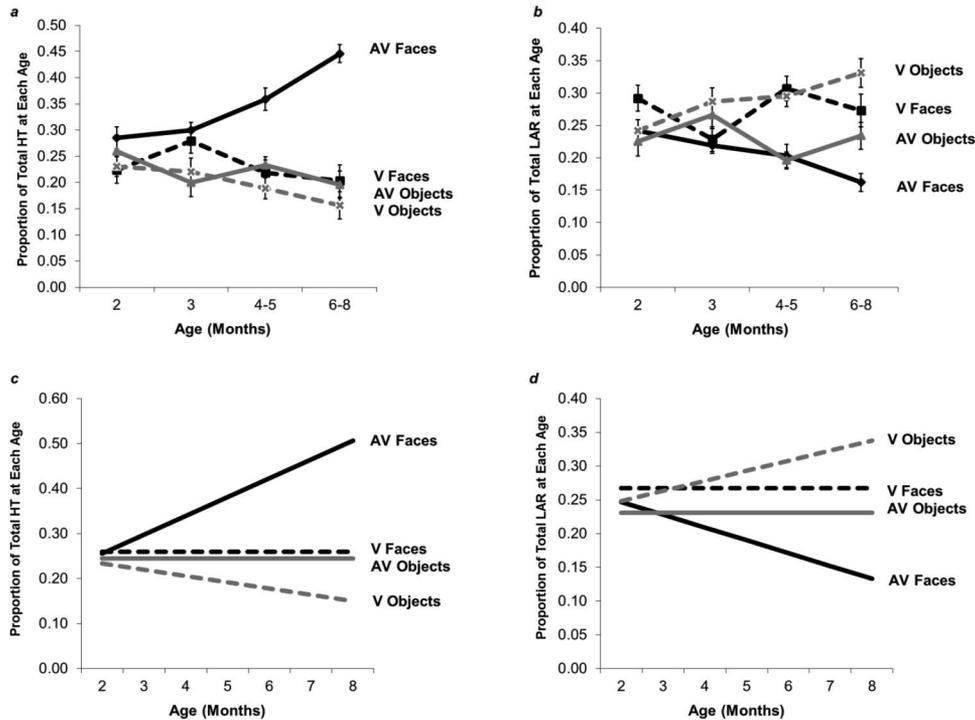
Although attention to faces and voices is considered foundational for the typical development of perception, cognition, lan-

guage, and social functioning, few studies have assessed when, if, or how attentional preferences for speaking faces emerge and change across infancy. The present study characterizes the developmental course of attention to audiovisual and visual faces and objects across early development. We created a unique and rich data set by combining and rescaling data from a large sample of 801 infants who had participated in infant-control habituation studies over the past two decades. This provides the first systematic picture of the development of attention maintenance according to two fundamental indices, HT and LAR, to dynamic visual and audiovisual events across 2 to 8 months of age. They serve as complementary indices of attention maintenance, with longer HT and lower LAR reflecting greater maintenance of attention.

Our data revealed several exciting new findings, as well as converging evidence for patterns of attention reported in the developmental literature. We found an overall decline in attention across 2 to 8 months of age (with decreasing HT and increasing LAR), consistent with the perspective that attention becomes more flexible and efficient across development (Colombo, 2001; Colombo, Shaddy, Richman, Maikranz, & Blaga, 2004; Courage et al., 2006; Ruff & Rothbart, 1996). Our results also indicated that HT and LAR become increasingly correlated across age, with shorter HT and more frequent looks away emerging across age. These findings indicate faster processing and greater control of attention across age, and a tighter coupling between these processes emerging across age. However, the overall decline in HT and increase in looking away across age did not hold for all event types. Rather, the patterns of attention to face versus object events,

<sup>5</sup> Slopes for proportion scores were also assessed for quadratic or cubic components. Analyses indicated only a significant quadratic function for one variable, audiovisual objects for LAR ( $p = .05$ ), with a decrease followed by a plateau or increase beyond 5 months. However, the difference between linear versus quadratic models was virtually zero ( $\Delta R^2 < .01$ ), indicating no significant gain by using a quadratic model. All predicted values fell within the expected range (0 to 1), indicating no bias in standard errors as a result of using proportion scores.

<sup>6</sup> To compensate for possible violations of normality, regression analyses were also conducted using a bootstrap approach. Bootstrap analyses confirmed the results of our standard regression analyses, and all slopes and differences between slopes that were significant remained significant with the bootstrap approach.



**Figure 4.** Proportion of total attention at each age. Means and standard errors for habituation times (HTs) and look-away rate (LAR; Figures 4a and 4b, respectively), and best-fitting regression lines for HT and LAR (Figures 4c and 4d, respectively). Figures 4a and 4c depict HT, and Figures 4b and 4d depict LAR, to audiovisual faces (AV Faces), audiovisual objects (AV Objects), visual faces (V Faces), and visual objects (V Objects). Proportion scores were derived by calculating HT and LAR for each event type at each age with respect to total HT and LAR across all event types at each age.

and to audiovisual versus visual stimulation, differed from one another and changed across age in several important ways.

### Face Versus Object Events

First, infants showed an increasing attentional advantage across age for the faces of people speaking over objects impacting a surface, consistent with prior findings of social preferences (Courage et al., 2006; Reynolds et al., 2013). They showed longer HT and lower LAR to face than to object events, and this pattern emerged gradually across age. Notably, this social preference was not evident at 2 months of age. Rather, it emerged at 3 months and became more evident by 4 to 5 and 6 to 8 months. Across age, infants showed a dramatic increase in the difference in overall HT to face over object events, with only a 2% difference at 2 months and a 30% difference by 6 to 8 months. The slopes for attention to face versus object events diverged significantly across age for both HT and LAR. Infants maintained high levels of attention to face events across 2 to 8 months of age, whereas attention to object events declined more steeply. These results are consistent with the perspective that preferences for social events emerge gradually across age. However, this developmental change is carried by attention to audiovisual face events, as illustrated by interactions with type of stimulation (see Audiovisual Face Events Versus Other Event Types).

### Audiovisual Versus Visual Events

Second, infants showed greater attention to audiovisual than silent visual events overall (longer HT and lower LAR), with increasingly greater differences across age. Longer attention maintenance to audiovisual than unimodal visual events was already evident by 2 months of age (but not at 3 months) and was strongest at 4 to 5 and 6 to 8 months, with only a 9% difference at 2 months and a 28% difference by 6 to 8 months. Slopes for both HT and LAR diverged significantly across age, indicating a steep decline in attention to unimodal visual events, but less decline for audiovisual events. This pattern reveals the early attentional salience of audiovisual events compared with silent visual events by 2 months and increasing across age. Findings were also qualified by interactions with event type (faces vs. objects; see Audiovisual Face Events Versus Other Event Types).

### Audiovisual Face Events Versus Other Event Types

Third, a novel finding consistent with our predictions revealed an attentional advantage for audiovisual speaking faces relative to each of the other event types (audiovisual objects, visual faces, and visual objects) that emerged gradually across development. Two-month-olds showed no attentional advantage for audiovisual faces. By 3 months of age, infants appeared to be in transition, showing greater attention maintenance to audiovisual faces than to two of

**Table 5**  
*Proportion Scores: Results From Regression Analysis Assessing Slopes Across Age for the Two Measures of Attention (Habituation Time, Look-Away Rate) as a Function of Event Type (Faces, Objects) and Type of Stimulation (Audiovisual, Visual)*

Measure	<i>b</i> estimate	<i>SE</i>	<i>p</i> value	<i>b</i> *
<b>Habituation time</b>				
Overall	.00	.003	.99	.01
Faces	.03	.005	<.001	.31
Objects	-.01	.004	.02	-.10
Audiovisual	.009	.004	.053	.09
Visual	-.015	.005	.003	-.15
Audiovisual faces	.042	.006	<.001	.45
Audiovisual objects	-.007	.006	.23	-.07
Visual faces	-.013	.009	.17	-.13
Visual objects	-.014	.005	.01	-.14
<b>Look-away rate</b>				
Overall	-.002	.003	.54	-.02
Faces	-.014	.004	.001	-.19
Objects	.004	.003	.22	.06
Audiovisual	-.01	.003	.005	-.13
Visual	.01	.004	.002	.17
Audiovisual faces	-.019	.005	<.001	-.26
Audiovisual objects	-.004	.005	.41	-.05
Visual faces	.004	.008	.58	.06
Visual objects	.015	.005	.001	.20

*Note.* Proportion scores were derived by calculating habituation time (HT) and look-away rate (LAR) for each event type at each age with respect to total HT and LAR across all event types at each age. *b* estimate = unstandardized regression coefficient; *SE* = standard error of the unstandardized coefficient; *b*\* = standardized regression coefficient.

the other event types (audiovisual objects, visual objects) for HT only. However, by 4 to 5 months (for HT) and 6 to 8 months (for HT and LAR), infants showed greater attention maintenance to speaking faces than to each of the other event types (audiovisual objects, visual faces, visual objects). Thus, the attentional advantage for audiovisual faces speaking was not present at 2 months of age and emerged gradually, with greater total fixation time by 4 to 5 months and reduced LARs by 6 to 8 months of age.

Regression analyses revealed a developmental trajectory for attention to audiovisual speaking faces that was distinct from that of each of the other event types. Across age, attention to audiovisual face events remained flat. This contrasts with the typical finding in the literature of an overall decline in looking time across 2 to 8 months of age. However, consistent with the literature, there was a dramatic and significant decline in attention to visual faces, visual objects, and audiovisual objects across age characterized by increasingly shorter HTs and more frequent looking away. The slopes of these three events differed significantly from that of the audiovisual face events. Thus, the difference in looking to audiovisual faces versus each of the other event types became more apparent with age. Differences in HT between audiovisual faces and each of the other event types at 2 months (average of 9%) versus 6 to 8 months (average 41%) underwent a dramatic 4.5-fold increase. Slopes for LARs also indicated that infants maintained high attention to audiovisual speaking faces across 2 to 8 months with only a marginal change. In contrast, LAR to each of the other event types increased significantly across age.

Another way of conceptualizing changes in attention maintenance across age is by using proportion scores to reflect selective attention to audiovisual faces compared with each of the other event types. The proportion of attention allocated to the audiovisual face events at each age (as a function of the total attention to all event types at each age) increased dramatically across 2 to 8 months of age. Regression analyses on proportion scores revealed a clear increase in attention maintenance to audiovisual faces across age and a decrease or maintenance of attention for each of the other event types across age. The slope for attention to audiovisual faces differed significantly from that of each of the other event types. Given limited attentional resources, particularly in infancy, it is the *relative allocation* of attention to different event types (selective attention) that provides the foundation and perceptual input upon which more complex cognitive, social, and language skills are built (Bahrick & Lickliter, 2012, 2014).

Taken together, these findings demonstrate a gradually increasing attentional advantage for audiovisual stimulation from people heard speaking over other event types across infancy. This attentional advantage is a result of infants maintaining high levels of attention to the faces of people speaking across age during a period when attention to other event types declines across age. In other words, the proportion of attention allocated to speaking faces relative to that of other event types increases across 2 to 8 months of age. This highlights the emerging attentional salience of audiovisual person events across 2 to 8 months of age, a salience that is highly adaptive. Caretakers scaffold infants' social, affective, and language development in face-to-face interactions (Flom, Lee, & Muir, 2007; Jaffe et al., 2001; Rochat, 1999). Enhanced attention to audiovisual face events creates greater opportunities for processing important dimensions of stimulation, including audiovisual affective expressions (Flom & Bahrick, 2007; Walker-Andrews, 1997), joint attention (Flom et al., 2007; Mundy & Burnette, 2005), speech (Fernald, 1985; Gogate & Hollich, 2010), and for increased engagement in social interactions and dyadic synchrony (Harrist & Waugh, 2002; Jaffe et al., 2001).

This pattern of emerging enhanced attention to speaking faces is also consistent with the central role of intersensory redundancy (e.g., common rhythm, tempo, and intensity changes arising from synchronous sights and sounds) in bootstrapping perceptual development (Bahrick & Lickliter, 2000, 2012, 2014). We have proposed that social events provide an extraordinary amount of redundancy across face, voice, and gesture, and that the salience of intersensory redundancy in audiovisual speech fosters early attentional preferences for these events, and, in turn, a developmental cascade leading to critical advances in perceptual, cognitive, and social development (Bahrick, 2010; Bahrick & Lickliter, 2012; Bahrick & Todd, 2012). Synchronous faces and voices elicit greater attentional salience and deeper processing than silent dynamic faces or faces presented with asynchronous voices, according to ERP measures (Reynolds et al., 2014). The present findings of overall preferences for audiovisual over visual events, for face over object events, and of a gradually emerging attentional advantage for audiovisual faces of people speaking over each of the other event types are consistent with this perspective.

However, demonstrating the critical role of intersensory redundancy (face-voice synchrony) in the attentional advantage for speaking faces and voices over other event types would require comparisons with an asynchronous control condition. Because

the present study did not include such a condition, it cannot be confirmed that intersensory redundancy was the basis for the growing attentional salience of speaking faces and voices. Alternative interpretations are also possible. For example, faces and voices provide a greater amount of stimulation than faces alone and/or the presence of the voice itself (rather than its synchrony with the movements of the face) and could enhance attention to speaking faces. However, prior research using asynchronous control conditions has ruled out both of these alternatives as explanations (Bahrack et al., 2002; Bahrack & Lickliter, 2000; Flom & Bahrack, 2007; Gogate & Bahrack, 1998). In each of these studies, attention and learning about properties of stimulation (rhythm, tempo, affect, speech sound-object relations) was facilitated by synchronous, but not by asynchronous, audiovisual stimulation. Further, synchrony between faces and voices was found to elicit deeper processing and greater attentional salience than asynchronous or dynamic visual faces (Reynolds et al., 2014). Thus, although the pivotal role of synchrony in promoting attentional salience in infancy is well established (Bahrack & Lickliter, 2002, 2012, 2014; Lewkowicz, 2000), more definitively characterizing its role in the emergence of attention to naturalistic face-voice events will require additional research. Further, longitudinal studies and assessments of relations with cognitive, social, and language outcomes will be needed to reveal more about the basis and implications of the divergent patterns of selective attention to face versus object events across age. Given that behavioral measures such as looking time can reflect different levels and types of attentional engagement (Reynolds & Richards, 2008), physiological and neural measures such as heart rate and ERP will also be important for revealing more about the nature of underlying attentional processes.

It also follows that children with impaired multisensory processing would show impairments in directing and maintaining attention to social events. Given that these events are typically more variable and complex and characterized by heightened levels of intersensory redundancy, impairments would be exaggerated for social compared with nonsocial events. Accordingly, children with autism show both impaired social attention (Dawson et al., 2004) and atypical intersensory processing (for a review, see Bahrack & Todd, 2012). Even a slight disturbance in multisensory processing could have cascading effects across development, beginning with decreased attention to social events, particularly people speaking, and leading to decreased opportunities for engagement in joint attention, language, and typical social interactions—all areas of impairment in children with autism (Bahrack, 2010; Mundy & Burnette, 2005). Further research is needed to more directly assess the role of intersensory processing in the typical and atypical development of social attention.

Why does the proportion of attention allocated to speaking faces relative to that of other event types increase across 2 to 8 months of age? Are infants processing the speaking faces and voices less efficiently than other event types? Or, in contrast, are they processing more information or processing the information more deeply? We favor the latter explanation. If speaking faces and voices are more complex, variable, and provide more information (Adolphs, 2001, 2009; Dawson et al., 1998), as well as exaggerated intersensory redundancy, compared with other event types, then longer attention maintenance (longer looking time and lower LAR) likely reflects continued and/or deeper processing of this

information. Research indicates synchrony elicits deeper processing and greater attentional salience than unimodal or asynchronous stimulation from the same events (Bahrack et al., 2013; Reynolds et al., 2014). Future studies using heart rate (see Richards & Casey, 1991) and neural measures of attention (ERP; Reynolds et al., 2014; Reynolds, Courage, & Richards, 2010) will be needed to determine the nature of relations between attention maintenance (as indexed by looking time and LAR) and processing speed, depth, and efficiency of processing.

## Comparisons With Other Studies

Our findings of a gradually emerging attentional advantage for speaking faces over object events across infancy are consistent with those of prior studies indicating that infants look longer to complex social events than simple nonsocial events (e.g., Sesame Street vs. geometric patterns); that they show deeper, more sustained attention to these events as indexed by greater decreases in heart rate; and that after 6 months of age, infants continue to show high levels of attention to dynamic, complex social events, whereas attention to static, simple events or nonsocial events reaches a plateau or declines (Courage et al., 2006; Reynolds et al., 2013; Richards, 2010). However, the developmental changes found in our study differ in some respects from those found in these prior studies. For example, Reynolds et al. (2013) found a decrease in attention to a complex social event (Sesame Street), both silent visual and audiovisual, across 3 to 9 months of age, and Courage et al. (2006) found an increase in looking to silent social events from 6.5 to 12 months. These inconsistencies are likely because of differences in stimuli (social events depicting Sesame Street vs. speaking faces), methods, and measures (HT vs. length of longest look, or average look length). Because neither of these studies included audiovisual face events, however, it is difficult to draw meaningful comparisons with our findings. In the present study, we presented a variety of faces of people (mostly women) speaking and objects consisting primarily of versions of toy hammers tapping various rhythms (and in our secondary data set, single and complex objects striking a surface). Generalization to other object and social event types should be made with caution, but the patterns observed *across age* are unaffected by these limitations. Our findings that dynamic speaking faces capture and maintain early attention, whereas attention to object events and visual-only events declines, illustrate the attentional “holding power” of audiovisual face events.

The present findings also revealed greater overall attention to audiovisual than visual-only events across infancy. Although prior research indicates infants show earlier, deeper, and/or more efficient processing of information in audiovisual events (redundantly specified properties) than the same properties in visual-only events (Bahrack et al., 2002; Bahrack & Lickliter, 2000, 2012; Flom & Bahrack, 2007; Reynolds et al., 2014), the literature on attention maintenance to audiovisual versus visual-only events is mixed. Some studies have shown greater looking to synchronous audiovisual than visual events (Bahrack & Lickliter, 2004; Bahrack et al., 2010); others report mixed results, with differences at some ages but not others (Reynolds et al., 2013); and others report no differences (Bahrack et al., 2002, 2013; Bahrack & Lickliter, 2004). The large sample and inclusion of multiple ages and conditions in the

present study provides a more comprehensive picture of these effects than previously available.

The present findings also address the long-standing theoretical debate regarding the origins of infant “social preferences.” Although some investigators have proposed that infant preferences for faces and social events are built in or arise from innate processing mechanisms (Balas, 2010; Gergely & Watson, 1999; Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, et al., 1991), others have argued that they emerge through experience with social events and result from general processing skills (Goldstein et al., 2003; Kuhl, Tsao, & Liu, 2003; Mastropieri & Turkewitz, 1999; Sai, 2005; Schaal, Marlier, & Soussignan, 1998, 2000). The present findings of a gradually emerging attentional advantage for audiovisual face events over other event types support the latter perspective regarding the critical role of experience with social events. Moreover, they are inconsistent with the proposal of innate face-processing mechanisms, as there was no evidence of a “face preference” or “social preference” at 2 months of age. Instead, 2-month-olds showed equal interest in the face and object events, and an attentional advantage for audiovisual events (both faces and objects) over visual events (both faces and objects). Our findings indicate a progressive differentiation across age, from no preference for faces at 2 months, to a preference for faces over object events by 3 months, followed by a preference for audiovisual face events over all other event types by 4 to 5 and 6 to 8 months of age. These findings highlight the important role of infant experience with dynamic social events and the audiovisual stimulation they provide.

### Summary and Conclusions

In sum, this study presents a novel approach to assessing typical developmental trajectories of infant attention to audiovisual and visual face versus object events, using two fundamental looking-time measures in a single study across a relatively wide age range (2 to 8 months). It provides a rich, new database and a more comprehensive picture of the development of attention than previously available. Our analyses are based on complete habituation data from an unusually large sample of 801 infants tested under uniform habituation conditions. Further, our events were dynamic and audiovisual, in contrast with static or silent visual stimuli used in most prior studies, enhancing the relevance of our findings to natural, multimodal events. We also assessed two complementary measures of attention, HT and LAR, typically not studied together. Converging data across these two different measures provides a new and more comprehensive picture of the development of attention to face and object events. Although overall attention maintenance declined across 2 to 8 months of age, converging with general trends reported in the literature, this decline did not characterize looking to coordinated faces and voices of people speaking. Instead, infants maintained high levels of attention to faces of speaking people across 2 to 8 months of age. This translates to an increasing attentional advantage for speaking faces relative to other event types across infancy. These findings are consistent with the hypothesis that enhanced attention to social events relative to object events emerges gradually as a function of experience with the social world, and that intersensory redundancy, available in natural, audiovisual stimulation, bootstraps attention to audiovisual speech in early development.

### References

- Abelkop, B. S., & Frick, J. E. (2003). Cross-task stability in infant attention: New perspectives using the still-face procedure. *Infancy*, *4*, 567–588. [http://dx.doi.org/10.1207/S15327078IN0404\\_09](http://dx.doi.org/10.1207/S15327078IN0404_09)
- Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*, 231–239. [http://dx.doi.org/10.1016/S0959-4388\(00\)00202-6](http://dx.doi.org/10.1016/S0959-4388(00)00202-6)
- Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. <http://dx.doi.org/10.1146/annurev.psych.60.110707.163514>
- Bahrnick, L. E. (1992). Infants’ perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology*, *53*, 180–199. [http://dx.doi.org/10.1016/0022-0965\(92\)90048-B](http://dx.doi.org/10.1016/0022-0965(92)90048-B)
- Bahrnick, L. E. (1994). The development of infants’ sensitivity to arbitrary intermodal relations. *Ecological Psychology*, *6*, 111–123. [http://dx.doi.org/10.1207/s15326969eco0602\\_2](http://dx.doi.org/10.1207/s15326969eco0602_2)
- Bahrnick, L. E. (2001). Increasing specificity in perceptual development: Infants’ detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology*, *79*, 253–270. <http://dx.doi.org/10.1006/jecp.2000.2588>
- Bahrnick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In J. G. Bremner & T. D. Wachs (Eds.), *The Wiley-Blackwell handbook of infant development: Vol. 1. Basic research* (2nd ed., pp. 120–165). Malden, MA: Wiley-Blackwell. <http://dx.doi.org/10.1002/9781444327564.ch4>
- Bahrnick, L. E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology*, *41*, 352–363. <http://dx.doi.org/10.1002/dev.10049>
- Bahrnick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*, 190–201. <http://dx.doi.org/10.1037/0012-1649.36.2.190>
- Bahrnick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 153–187). San Diego, CA: Academic Press.
- Bahrnick, L. E., & Lickliter, R. (2004). Infants’ perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective & Behavioral Neuroscience*, *4*, 137–147. <http://dx.doi.org/10.3758/CABN.4.2.137>
- Bahrnick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183–206). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199586059.003.0008>
- Bahrnick, L. E., & Lickliter, R. (2014). Learning to attend selectively: The dual role of intersensory redundancy. *Current Directions in Psychological Science*, *23*, 414–420. <http://dx.doi.org/10.1177/0963721414549187>
- Bahrnick, L. E., Lickliter, R., & Castellanos, I. (2013). The development of face perception in infancy: Intersensory interference and unimodal visual facilitation. *Developmental Psychology*, *49*, 1919–1930. <http://dx.doi.org/10.1037/a0031238>
- Bahrnick, L. E., Lickliter, R., Castellanos, I., & Todd, J. T. (2015). Intersensory redundancy facilitates infant detection of tempo: Extending predictions of the intersensory redundancy hypothesis. *Infancy*, *20*, 377–404. <http://dx.doi.org/10.1111/inf.12081>
- Bahrnick, L. E., Lickliter, R., Castellanos, I., & Vaillant-Molina, M. (2010). Increasing task difficulty enhances effects of intersensory redundancy: Testing a new prediction of the intersensory redundancy hypothesis. *Developmental Science*, *13*, 731–737. <http://dx.doi.org/10.1111/j.1467-7687.2009.00928.x>

- Bahrack, L. E., Lickliter, R., & Flom, R. (2006). Up versus down: The role of intersensory redundancy in the development of infants' sensitivity to the orientation of moving objects. *Infancy*, 9, 73–96. [http://dx.doi.org/10.1207/s15327078in0901\\_4](http://dx.doi.org/10.1207/s15327078in0901_4)
- Bahrack, L. E., Lickliter, R., Shuman, M., Batista, L. C., Castellanos, I., & Newell, L. C. (2005, November). *The development of infant voice discrimination: From unimodal auditory to bimodal audiovisual stimulation*. Poster presented at the International Society of Developmental Psychobiology, Washington, DC.
- Bahrack, L. E., Lickliter, R., Shuman, M., Batista, L. C., & Grandez, C. (2003, April). *Infant discrimination of voices: Predictions of the intersensory redundancy hypothesis*. Poster presented at the Society for Research in Child Development, Tampa, FL.
- Bahrack, L. E., Newell, L. C., Shuman, M., & Ben, Y. (2007, March). *Three-month-old infants recognize faces in unimodal visual but not bimodal audiovisual stimulation*. Poster presented at the Society for Research in Child Development, Boston, MA.
- Bahrack, L. E., Shuman, M. A., & Castellanos, I. (2008, March). *Face-voice synchrony directs selective listening in four-month-old infants*. Poster presented at the International Conference on Infant Studies, Vancouver, Canada.
- Bahrack, L. E., & Todd, J. T. (2012). Multisensory processing in autism spectrum disorders: Intersensory processing disturbance as a basis for atypical development. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 1453–1508). Cambridge, MA: MIT Press.
- Balas, B. (2010). Using innate visual biases to guide face learning in natural scenes: A computational investigation. *Developmental Science*, 13, 469–478. <http://dx.doi.org/10.1111/j.1467-7687.2009.00901.x>
- Bornstein, M. H., & Colombo, J. (2012). Infant cognitive functioning and mental development. In S. M. Pauen (Ed.), *Early childhood development and later achievement* (pp. 118–147). Cambridge, UK: Cambridge University Press.
- Castellanos, I., & Bahrack, L. E. (2007, March). *Intersensory redundancy educates infants' attention to amodal properties in speech during early development*. Poster presented at the Society for Research in Child Development, Boston, MA.
- Castellanos, I., Shuman, M., & Bahrack, L. E. (2004, May). *Intersensory redundancy facilitates infants' perception of meaning in speech passages*. Poster presented at the International Conference on Infant Studies, Chicago, IL.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, 52, 337–367. <http://dx.doi.org/10.1146/annurev.psych.52.1.337>
- Colombo, J. (2004). Visual attention in infancy: Process and product in early cognitive development. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 329–341). New York, NY: Guilford Press.
- Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 283–322). San Diego, CA: Academic Press. [http://dx.doi.org/10.1016/S0065-2407\(06\)80010-8](http://dx.doi.org/10.1016/S0065-2407(06)80010-8)
- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, 92, 225–234. <http://dx.doi.org/10.1016/j.nlm.2008.06.002>
- Colombo, J., Shaddy, D. J., Richman, W. A., Maikranz, J. M., & Blaga, O. M. (2004). The developmental course of habituation in infancy and preschool outcome. *Infancy*, 5, 1–38. [http://dx.doi.org/10.1207/s15327078in0501\\_1](http://dx.doi.org/10.1207/s15327078in0501_1)
- Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, 77, 680–695. <http://dx.doi.org/10.1111/j.1467-8624.2006.00897.x>
- Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479–485. <http://dx.doi.org/10.1023/A:1026043926488>
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40, 271–283. <http://dx.doi.org/10.1037/0012-1649.40.2.271>
- Doheny, L., Hurwitz, S., Insoft, R., Ringer, S., & Lahav, A. (2012). Exposure to biological maternal sounds improves cardiorespiratory regulation in extremely preterm infants. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25, 1591–1594. <http://dx.doi.org/10.3109/14767058.2011.648237>
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 102, 17245–17250. <http://dx.doi.org/10.1073/pnas.0502205102>
- Feldman, R. (2007). Parent-infant synchrony: Biological foundations and developmental outcomes. *Current Directions in Psychological Science*, 16, 340–345. <http://dx.doi.org/10.1111/j.1467-8721.2007.00532.x>
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, 8, 181–195. [http://dx.doi.org/10.1016/S0163-6383\(85\)80005-9](http://dx.doi.org/10.1016/S0163-6383(85)80005-9)
- Ferrara, C., & Hill, S. D. (1980). The responsiveness of autistic children to the predictability of social and nonsocial toys. *Journal of Autism and Developmental Disorders*, 10, 51–57. <http://dx.doi.org/10.1007/BF02408432>
- Flom, R., & Bahrack, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology*, 43, 238–252. <http://dx.doi.org/10.1037/0012-1649.43.1.238>
- Flom, R., Lee, K., & Muir, D. (2007). *Gaze-following: Its development and significance*. Mahwah, NJ: Erlbaum.
- Gergely, G., & Watson, J. S. (1999). Early socio-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 101–136). Mahwah, NJ: Erlbaum.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. East Norwalk, CT: Appleton-Century-Crofts.
- Gogate, L. J., & Bahrack, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, 69, 133–149. <http://dx.doi.org/10.1006/jecp.1998.2438>
- Gogate, L. J., Bahrack, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71, 878–894. <http://dx.doi.org/10.1111/1467-8624.00197>
- Gogate, L. J., & Hollich, G. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review*, 117, 496–516. <http://dx.doi.org/10.1037/a0019049>
- Gogate, L., Maganti, M., & Bahrack, L. E. (2015). Cross-cultural evidence for multimodal motherese: Asian Indian mothers' adaptive use of synchronous words and gestures. *Journal of Experimental Child Psychology*, 129, 110–126. <http://dx.doi.org/10.1016/j.jecp.2014.09.002>
- Gogate, L. J., Walker-Andrews, A. S., & Bahrack, L. E. (2001). The intersensory origins of word comprehension: An ecological-dynamic systems view. *Developmental Science*, 4, 1–18. <http://dx.doi.org/10.1111/1467-7687.00143>
- Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 100, 8030–8035. <http://dx.doi.org/10.1073/pnas.1332441100>
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56, 544–549.

- Greenough, W. T., & Black, J. E. (1992). Induction of brain structure by experience: Substrates for cognitive development. In M. R. Gunnar and C. A. Nelson (Eds.), *Developmental behavioral neuroscience: Minnesota Symposium on Child Psychology* (Vol. 24, pp. 155–200). Hillsdale, NJ: Erlbaum.
- Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental Review*, 22, 555–592. [http://dx.doi.org/10.1016/S0273-2297\(02\)00500-2](http://dx.doi.org/10.1016/S0273-2297(02)00500-2)
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- Hunnis, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, 6, 231–255. [http://dx.doi.org/10.1207/s15327078in0602\\_5](http://dx.doi.org/10.1207/s15327078in0602_5)
- Jaccard, J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child and Adolescent Psychology*, 31, 130–146. [http://dx.doi.org/10.1207/S15374424JCCP3101\\_15](http://dx.doi.org/10.1207/S15374424JCCP3101_15)
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, 66(2), 1–131.
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40, 1–19. [http://dx.doi.org/10.1016/0010-0277\(91\)90045-6](http://dx.doi.org/10.1016/0010-0277(91)90045-6)
- Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, 3, 335–344. <http://dx.doi.org/10.1162/jocn.1991.3.4.335>
- Kavšek, M. (2013). The comparator model of infant visual habituation and dishabituation: Recent insights. *Developmental Psychobiology*, 55, 793–808. <http://dx.doi.org/10.1002/dev.21081>
- Kim, H. I., & Johnson, S. P. (2014). Detecting “infant-directedness” in face and voice. *Developmental Science*, 17, 621–627. <http://dx.doi.org/10.1111/desc.12146>
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16, 1412–1425. <http://dx.doi.org/10.1162/0898929042304796>
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 100, 9096–9101. <http://dx.doi.org/10.1073/pnas.1532872100>
- Legerstee, M., Pomerleau, A., Malcuit, G., & Feider, H. (1987). The development of infants' responses to people and a doll: Implications for research in communication. *Infant Behavior & Development*, 10, 81–95. [http://dx.doi.org/10.1016/0163-6383\(87\)90008-7](http://dx.doi.org/10.1016/0163-6383(87)90008-7)
- Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, 126, 281–308. <http://dx.doi.org/10.1037/0033-2909.126.2.281>
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*, 35, 204–214. [http://dx.doi.org/10.1002/\(SICI\)1098-2302\(199911\)35:3<204::AID-DEV5>3.0.CO;2-V](http://dx.doi.org/10.1002/(SICI)1098-2302(199911)35:3<204::AID-DEV5>3.0.CO;2-V)
- Maurer, D., & Barrera, M. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development*, 52, 196–202. <http://dx.doi.org/10.2307/1129230>
- Mundy, P., & Burnette, C. (2005). Joint attention and neurodevelopmental models of autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders, Vol. 1: Diagnosis, development, neurobiology, and behavior* (3rd ed., pp. 650–681). Hoboken, NJ: Wiley. <http://dx.doi.org/10.1002/9780470939345.ch25>
- Newell, L. C., Castellanos, I., Grossman, R., & Bahrack, L. E. (2009, March). *Bimodal, synchronous displays, but not unimodal, visual displays, elicit gender discrimination in 6-month-old infants*. Poster presented at the Society for Research in Child Development, Denver, CO.
- Oakes, L. M., Kannass, K. N., & Shaddy, D. J. (2002). Developmental changes in endogenous control of attention: The role of target familiarity on infants' distraction latency. *Child Development*, 73, 1644–1655. <http://dx.doi.org/10.1111/1467-8624.00496>
- Pérez-Edgar, K., McDermott, J. N. M., Korelitz, K., Degnan, K. A., Curby, T. W., Pine, D. S., & Fox, N. A. (2010). Patterns of sustained attention in infancy shape the developmental trajectory of social behavior from toddlerhood through adolescence. *Developmental Psychology*, 46, 1723–1730. <http://dx.doi.org/10.1037/a0021064>
- Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. *Developmental Psychology*, 48, 827–835. <http://dx.doi.org/10.1037/a0025530>
- Reynolds, G. D., Bahrack, L. E., Lickliter, R., & Guy, M. W. (2014). Neural correlates of intersensory processing in 5-month-old infants. *Developmental Psychobiology*, 56, 355–372. <http://dx.doi.org/10.1002/dev.21104>
- Reynolds, G. D., Courage, M. L., & Richards, J. E. (2010). Infant attention and visual preferences: Converging evidence from behavior, event-related potentials, and cortical source localization. *Developmental Psychology*, 46, 886–904. <http://dx.doi.org/10.1037/a0019670>
- Reynolds, G. D., & Richards, J. E. (2008). Infant heart rate: A developmental psychophysiological perspective. In L. A. Schmidt & S. J. Segalowitz (Eds.), *Developmental psychophysiology: Theory, systems, and methods* (pp. 173–212). Cambridge, MA: Cambridge University Press.
- Reynolds, G. D., Zhang, D., & Guy, M. W. (2013). Infant attention to dynamic audiovisual stimuli: Look duration from 3 to 9 months of age. *Infancy*, 18, 554–577. <http://dx.doi.org/10.1111/j.1532-7078.2012.00134.x>
- Richards, J. E. (2010). The development of attention to simple and complex visual stimuli in infants: Behavioral and psychophysiological measures. *Developmental Review*, 30, 203–219. <http://dx.doi.org/10.1016/j.dr.2010.03.005>
- Richards, J. E., & Casey, B. J. (1991). Heart rate variability during attention phases in young infants. *Psychophysiology*, 28, 43–53. <http://dx.doi.org/10.1111/j.1469-8986.1991.tb03385.x>
- Richards, J. E., & Turner, E. D. (2001). Extended visual fixation and distractibility in children from six to twenty-four months of age. *Child Development*, 72, 963–972. <http://dx.doi.org/10.1111/1467-8624.00328>
- Rochat, P. (1999). *Early social cognition: Understanding others in the first months of life*. Mahwah, NJ: Erlbaum.
- Rochat, P. (2007). Intentional action arises from early reciprocal exchanges. *Acta Psychologica*, 124, 8–25. <http://dx.doi.org/10.1016/j.actpsy.2006.09.004>
- Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2005). Pathways from prematurity and infant abilities to later cognition. *Child Development*, 76, 1172–1184. <http://dx.doi.org/10.1111/j.1467-8624.2005.00842.x-i1>
- Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. New York, NY: Oxford University Press.
- Sai, F. Z. (2005). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development*, 14, 29–50. <http://dx.doi.org/10.1002/icd.376>
- Schaal, B., Marlier, L., & Soussignan, R. (1998). Olfactory function in the human fetus: Evidence from selective neonatal responsiveness to the odor of amniotic fluid. *Behavioral Neuroscience*, 112, 1438–1449. <http://dx.doi.org/10.1037/0735-7044.112.6.1438>
- Schaal, B., Marlier, L., & Soussignan, R. (2000). Human foetuses learn odours from their pregnant mother's diet. *Chemical Senses*, 25, 729–737. <http://dx.doi.org/10.1093/chemse/25.6.729>

- Shaddy, D. J., & Colombo, J. (2004). Developmental changes in infant attention to dynamic and static stimuli. *Infancy*, 5, 355–365. [http://dx.doi.org/10.1207/s15327078in0503\\_6](http://dx.doi.org/10.1207/s15327078in0503_6)
- Sheese, B. E., Rothbart, M. K., Posner, M. I., White, L. K., & Fraundorf, S. H. (2008). Executive attention and self-regulation in infancy. *Infant Behavior & Development*, 31, 501–510. <http://dx.doi.org/10.1016/j.infbeh.2008.02.001>
- Simion, F., Cassia, V. M., Turati, C., & Valenza, E. (2001). The origins of face perception: Specific versus non-specific mechanisms. *Infant and Child Development*, 10, 59–65. <http://dx.doi.org/10.1002/icd.247>
- Smith, N. A., & Strader, H. L. (2014). Infant-directed visual prosody: Mothers' head movements and speech acoustics. *Interaction Studies*, 15, 38–54. <http://dx.doi.org/10.1075/is.15.1.02smi>
- Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. New York, NY: Basic Books.
- Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 892–903. <http://dx.doi.org/10.1037/0096-1523.22.4.892>
- Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121, 437–456. <http://dx.doi.org/10.1037/0033-2909.121.3.437>
- Watson, T. L., Robbins, R. A., & Best, C. T. (2014). Infant perceptual development for faces and spoken words: An integrated approach. *Developmental Psychobiology*, 56, 1454–1481. <http://dx.doi.org/10.1002/dev.21243>

Received December 1, 2014

Revision received February 5, 2016

Accepted June 3, 2016 ■

### Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at [Reviewers@apa.org](mailto:Reviewers@apa.org). Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit <http://www.apa.org/pubs/authors/review-manuscript-ce-video.aspx>.