

The Development of Face Perception in Infancy: Intersensory Interference and Unimodal Visual Facilitation

Lorraine E. Bahrick, Robert Lickliter, and Irina Castellanos
Florida International University

Although research has demonstrated impressive face perception skills of young infants, little attention has focused on conditions that enhance versus impair infant face perception. The present studies tested the prediction, generated from the intersensory redundancy hypothesis (IRH), that face discrimination, which relies on detection of visual featural information, would be impaired in the context of intersensory redundancy provided by audiovisual speech and enhanced in the absence of intersensory redundancy (unimodal visual and asynchronous audiovisual speech) in early development. Later in development, following improvements in attention, faces should be discriminated in both redundant audiovisual and nonredundant stimulation. Results supported these predictions. Two-month-old infants discriminated a novel face in unimodal visual and asynchronous audiovisual speech but not in synchronous audiovisual speech. By 3 months, face discrimination was evident even during synchronous audiovisual speech. These findings indicate that infant face perception is enhanced and emerges developmentally earlier following unimodal visual than synchronous audiovisual exposure and that intersensory redundancy generated by naturalistic audiovisual speech can interfere with face processing.

Keywords: audiovisual event perception, face–voice processing, face discrimination

The faces of people are highly salient to young infants and convey information critical for social interaction, including emotion, intention, and individual identity. A large body of research has demonstrated remarkable face perception skills of very young infants (e.g., Bartrip, Morton, & de Schonen, 2001; Cassia, Turati, & Simion, 2004; Johnson, Dziurawiec, Ellis, & Morton, 1991; Mondloch et al., 1999). For example, shortly following birth, infants can discriminate, prefer, and show memory for their mother's face over the face of a stranger in silent visual displays (Bushnell, 2001; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995; Sai, 2005), and they can do so after 1 month of age even when external features have been masked (Bartrip et al., 2001). By 2–4 months, infants discriminate between the faces of unfamiliar adults. Three-month-olds discriminate between photographs of adults both within and outside their own racial group (Kelly et al., 2009). Two-

and three-month-olds discriminate between the faces of two unfamiliar women in live, static poses both with and without external cues masked (Blass & Camp, 2004). At 4 and 6 months (but not at 2 months), infants can discriminate the moving faces of two unfamiliar women even in the context of synchronous audiovisual speech and can subsequently match their particular faces with their voices (Bahrick, Hernandez-Reif, & Flom, 2005). Although 2-month-olds showed no evidence of face–voice matching, they were able to discriminate between the two faces in a silent, dynamic, visual control condition. Some researchers have proposed that faces represent a “special” class of stimulation and that that they are innately preferred over (Goren, Sarty, & Wu, 1975; Johnson & Morton, 1991; Mondloch et al., 1999) and processed differently from other stimuli (Tanaka & Sengco, 1997; Thompson & Massaro, 1989; Ward, 1989; Yovel & Duchaine, 2006). In contrast, others have suggested that face processing is governed by the same principles as object perception and that faces become salient as a result of experience interacting with people and the development of expertise with this particular class of stimulation (e.g., Bahrick & Newell, 2008; Gauthier & Nelson, 2001; Nelson, 2003).

The majority of research on face perception has focused on infant perception of static silent faces, even though faces are typically perceived in the context of individuals engaged in dynamic, multimodal events such as everyday activities and audiovisual speech. Despite findings that infants are adept at perceiving dynamic, multimodal events (Bahrick, 2010; Lickliter & Bahrick, 2004; Lewkowicz, 2000; Lewkowicz & Lickliter, 1994), and at perceiving information in dynamic, speaking faces (Bahrick et al., 2005; Bahrick, Moss, & Fadil, 1996; Hollich, Newman, & Juszyk, 2005; Lewkowicz, 2010; Walker-Andrews, 1997), infants have poor face perception skills when faces are seen in the context of naturalistic everyday activities (Bahrick, Gogate, & Ruiz, 2002;

This article was published Online First December 17, 2012.

Lorraine E. Bahrick, Robert Lickliter, and Irina Castellanos, Department of Psychology, Florida International University.

This research was supported by National Institute of Mental Health Grant R01 MH62226 and National Institute of Child Health and Human Development Grants K02 HD064943 and RO1 HD053776 awarded to Lorraine Bahrick and by National Science Foundation Grant BCS1057898 awarded to Robert Lickliter. Irina Castellanos was supported by National Institutes of Health/National Institute of General Medical Sciences Grant R25 GM061347.

Portions of these data were reported at the International Conference on Infant Studies in 2004 and 2006 and at the International Multisensory Research Forum in 2004.

Correspondence concerning this article should be addressed to Lorraine E. Bahrick, Department of Psychology, Florida International University, Miami, FL 33199. E-mail: bahrick@fiu.edu

Bahrack & Newell, 2008). For example, 5-month-olds show discrimination and long-term memory (across a 7-week delay) for repetitive activities (e.g., brushing teeth, brushing hair, blowing bubbles) but no discrimination or memory for the faces of the women engaged in the activities. Infants discriminated the faces of the women engaged in the activities only after twice the exposure time required for discriminating the actions (320 s vs. 160 s) or after the shorter exposure time of 160 s when movement was eliminated and the faces were static. These effects have yet to be investigated in children and adults.

These findings demonstrate the superiority of action perception over face perception during early development and reflect the effects of early stimulus-driven attentional salience hierarchies in infancy. Actions are more salient than faces, and they receive earlier and greater attentional selectivity than do faces in the context of actions (Bahrack, Gogate, & Ruiz, 2002; Bahrack & Newell, 2008). Attention is first drawn to motion at the expense of the appearance of the object engaged in motion. Once the salient motion is processed, attention progresses to lower levels of the salience hierarchy, including the moving object and its distinctive features. Further evidence for the critical role of attentional salience comes from a control study demonstrating that infants were not impaired at discriminating faces in the context of actions. After infants were habituated to the actions, they were then able to attend to the faces of the women engaged in the actions and discriminated among them (Bahrack & Newell, 2008). Thus, the superiority of action over face perception was not due to impaired discrimination of dynamic faces; it was a result of attentional allocation. In fact, perception of faces is facilitated in dynamic as compared with static displays and provides invariant information for facial features across changes over time (Bahrack et al., 1996; Otsuka et al., 2009). The operation of salience hierarchies in attention allocation is especially evident when attentional resources are most limited (Bahrack, 2010; see also Adler, Gerhardstein, & Rovee-Collier, 1998), as is the case during infancy (see also Craik & Byrd, 1982, and Craik, Luo, & Sakuta, 2010, for examples with adults and the elderly).

Another critical stimulus factor that organizes early attentional salience hierarchies and that should have a profound effect on face perception is *intersensory redundancy*. Intersensory redundancy is the synchronous co-occurrence of the same *amodal information* (i.e., temporal, spatial, or intensity patterns) across multiple sense modalities (e.g., the rhythm and tempo common to synchronous audible and visible speech; Bahrack & Lickliter, 2000, 2003, 2012). Detection of redundancy across the senses gives rise to salience hierarchies that organize and guide selective attention and perceptual learning in early development. Detection of intersensory redundancy enables naïve perceivers to determine which sights and sounds constitute unitary audiovisual events, such as a speaking person (Bahrack & Lickliter, 2003, 2012; Lewkowicz, 2000, 2010), and in turn, to detect nested amodal properties specifying aspects such as affect and the prosody of speech (Bahrack, 2001, 2010; Castellanos & Bahrack, 2008; Castellanos, Shuman, & Bahrack, 2004; Flom & Bahrack, 2007; Vaillant-Molina & Bahrack, 2012). Because of limited attentional resources, particularly in early development, and the vast amount of available sensory stimulation, selective attention to certain properties of stimulation always occurs at the expense of attention to other properties.

The intersensory redundancy hypothesis (IRH; Bahrack, 2010; Bahrack & Lickliter, 2000, 2003, 2012), a model of the development of early attentional selectivity, describes how redundancy across the senses organizes early attentional salience hierarchies and provides the foundation for early perceptual development. According to the IRH, information presented redundantly and in temporal synchrony across two or more senses (e.g., amodal information such as rhythm, tempo, and the prosody of speech) is highly salient. It recruits selective attention and facilitates perceptual learning of amodal properties at the expense of less salient modality-specific properties (e.g., visual information, such as features of the face, color, and pattern, or auditory information, such as timbre and pitch; Bahrack, 2010; Bahrack & Lickliter, 2000, 2003, 2012; Bahrack, Lickliter, & Flom, 2004). Detection of affect and the prosody of speech is based on perception of amodal information because they are comprised of distinctive patterns of tempo, rhythm, and intensity changes that are redundant across visual (face) and auditory (voice) stimulation. A large body of research has demonstrated that infants are adept perceivers of intersensory redundancy and that the salience of redundancy guides early attentional selectivity in both human and nonhuman animal infants (e.g., Bahrack, Flom, & Lickliter, 2002; Bahrack & Lickliter, 2000, 2003; Lewkowicz, 2004; Lickliter, Bahrack, & Honeycutt, 2002, 2004; Lickliter, Bahrack, & Markham, 2006). Intersensory redundancy thus exerts a powerful organizing force on the direction of perceptual development by focusing attention on amodal, redundantly specified properties at the expense of nonredundantly specified, modality-specific properties of events during early development.

The IRH has proposed and generated evidence for two important principles of early event perception, *intersensory facilitation* and *unimodal facilitation*. Intersensory facilitation characterizes perception of naturalistic, multimodal events such as an object striking a surface or audiovisual speech. Intersensory facilitation refers to the principle that redundantly specified amodal properties (rhythm, tempo, intensity, etc.) are detected more easily and earlier in development when they are perceived in bimodal (or multimodal) synchronous stimulation than when the same amodal properties are nonredundantly specified in unimodal stimulation. Perception of amodal properties, such as temporal synchrony (Bahrack, 1988, 2001; Lewkowicz, 2000, 2010), rhythm (Bahrack & Lickliter, 2000), tempo, intensity, prosody, and affect (Bahrack, Flom, & Lickliter, 2002; Bahrack, Lickliter, Castellanos, & Vaillant-Molina, 2010; Castellanos, 2007; Castellanos & Bahrack, 2008; Castellanos et al., 2004; Castellanos, Vaillant-Molina, Lickliter, & Bahrack, 2006; Flom & Bahrack, 2007), is enhanced in bimodal audiovisual as compared with unimodal stimulation. For example, at 4 months, infants can discriminate affect in audiovisual speech but not in unimodal auditory or visual speech (Flom & Bahrack, 2007), and quail chicks learn and remember a maternal call following synchronous audiovisual exposure but not following the equivalent amount of unimodal auditory or asynchronous audiovisual exposure (Lickliter et al., 2002, 2004). Intersensory facilitation is now a well-documented principle of early perceptual development (see Bahrack & Lickliter, 2000; Farzin, Charles, & Rivera, 2009; Frank, Slemmer, Marcus, & Johnson, 2009; Gogate & Hollich, 2010; Gogate, Walker-Andrews & Bahrack, 2001; Hollich et al., 2005; Jordan, Suanda, & Brannon, 2008; Lewkowicz, 2004).

In contrast, unimodal facilitation is evident when stimulation is detectable through only one sense modality (unimodal) and no intersensory redundancy is available to direct attention to amodal properties of events. In unimodal stimulation (e.g., speaking on the phone; viewing a silent face or action), information is available to only one sensory modality at a time. In this case, the IRH proposes that selective attention and learning are promoted to modality-specific properties of stimulation (properties that can be specified only through a particular sensory modality) at the expense of amodal properties (Bahrick, 2010; Bahrick & Lickliter, 2000, 2003, 2012; Bahrick et al., 2004). *Unimodal facilitation* refers to the principle that nonredundantly specified modality-specific properties (such as color, visual pattern, pitch, and timbre) are detected more easily and earlier in development when they are perceived in unimodal stimulation than when the same properties are perceived in redundant bimodal (or multimodal) stimulation, because there is no attentional competition from salient intersensory redundancy. The principles of unimodal and intersensory facilitation are particularly evident in early development because attentional resources are most limited, and thus attention progresses slowly along the attentional salience hierarchy. Later in development, as infant attention becomes more flexible and efficient, infants can discriminate amodal and modality-specific properties in both unimodal and multimodal stimulation (Bahrick, 2010; Bahrick & Lickliter, 2004, 2012).

In contrast with the large body of research on the facilitating role of intersensory redundancy for perceptual processing, only a few studies have focused on unimodal facilitation and the interfering role of intersensory redundancy for perceptual processing. These studies have demonstrated that unimodal stimulation facilitates perceptual processing of modality-specific properties of events. Perception of modality-specific properties such as orientation (Bahrick, Lickliter, & Flom, 2006; Flom & Bahrick, 2010), color/pattern (Vaillant-Molina, Gutierrez, & Bahrick, 2005), and pitch/timbre (Bahrick, Lickliter, Shuman, Batista, & Grande, 2003; Vaillant, Bahrick, & Lickliter, 2012) is enhanced in unimodal as compared with bimodal audiovisual stimulation. For example, 5- and 9-month-old infants discriminated and showed long-term memory for a change in the orientation of a toy hammer tapping (upward against a ceiling vs. downward against a floor) when they could see the hammer tapping (unimodal visual) but not when they could see and hear the natural synchronous audiovisual stimulation together (Bahrick et al., 2006; Flom & Bahrick, 2010). The audiovisual condition provided intersensory redundancy, which attracts attention to redundantly specified amodal properties, such as the rhythm and tempo of the hammer (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002), and interferes with attention to visual information, such as the direction of motion or orientation of the hammer. It was not until 9 months of age that infants could detect and remember the orientation of the hammer in the presence of intersensory redundancy. Performance in an asynchronous control condition confirmed the interfering role of intersensory redundancy. In asynchronous stimulation, intersensory redundancy is eliminated, but the overall amount and type of stimulation are equated with that of audiovisual synchronous stimulation from the same event. As predicted, instead of impairing perception of orientation, the asynchronous soundtrack enhanced infant perception of orientation when compared with the synchronized soundtrack (Bahrick et al., 2006). This unimodal

facilitation illustrates enhanced perception of modality-specific properties in unimodal as compared with multimodal stimulation as a result of the lack of attentional competition from intersensory redundancy (see Bahrick & Lickliter, 2012, for further discussion).

Importantly, faces are typically encountered as part of dynamic, multimodal speech events. However, little is known about infant perception of dynamic speaking faces or about the conditions that facilitate or attenuate face perception in early development. We propose that infant face perception, like infant perception of non-social object events, is guided by general principles of perceptual development and the salience of intersensory redundancy. The principles of intersensory and unimodal facilitation proposed by the IRH provide the basis for strong predictions regarding conditions that promote versus attenuate face processing in early development. Discriminating among faces requires detection of modality-specific information, including visual pattern, color, facial features, and their spatial configuration (Cohen & Cashion, 2001; Rotshtein, Geng, Driver, & Dolan, 2007; Tanaka & Sengco, 1997). Therefore, in early development, infants should show unimodal facilitation of face perception, such that face discrimination is enhanced in unimodal visual stimulation (in the absence of synchronous vocal stimulation) and attenuated in synchronous bimodal stimulation (audiovisual speech), where salient redundant amodal information competes for attention. Predictions of the IRH hold for tasks that are relatively novel or difficult in relation to the skills of the perceiver (Bahrick & Lickliter, 2000, 2003, 2012; Bahrick et al., 2010), and thus these effects should be most apparent for perception of unfamiliar faces. For familiar faces, such as that of the mother, recognition would likely be based on a variety of factors. Further, given the infant's extended exposure to the face of the mother in both silent and audiovisual conditions, perceptual differentiation of modality-specific visual properties specifying facial features and their arrangement should have progressed, allowing for flexible identification across a variety of conditions, including contexts that provide intersensory redundancy and those that do not.

Research has already demonstrated that redundantly specified properties conveyed by faces (e.g., affect, prosody of speech) are perceived earlier in development and detected more readily in synchronous bimodal audiovisual stimulation than in unimodal stimulation (intersensory facilitation; Castellanos & Bahrick, 2008; Castellanos et al., 2004; Flom & Bahrick, 2007; Vaillant-Molina & Bahrick, 2012). However, it is not known if perception of nonredundantly specified properties conveyed by faces (i.e., those that permit individual identification based on facial features and their configuration) is facilitated in unimodal as compared with bimodal audiovisual stimulation (i.e., unimodal facilitation). If face perception is governed by general principles of event perception that apply to other objects, unimodal facilitation should be evident in the domain of face perception. Given that intersensory redundancy captures selective attention and thus interferes with perception of modality-specific information in a variety of domains in early development, intersensory redundancy should impair perception and discrimination of faces as well. If so, this would provide evidence that face perception illustrates general principles of perceptual learning and development and should not be considered a special class of stimulation in the sense that its attentional salience is governed by processes that differ from those of other objects and events. In the experiments reported here, we

thus tested the principle of unimodal facilitation by assessing the development of infants' discrimination of faces in redundant and nonredundant stimulation.

Given that discrimination of faces is based on detection of facial features and their configuration, information specific to visual stimulation (Cohen & Cashon, 2001; Rotshtein et al., 2007; Tanaka & Sengco, 1997), we predicted that for very young infants presented with the faces of unfamiliar women speaking, face discrimination would be enhanced in unimodal visual speech (i.e., silent speech), where no intersensory redundancy is present, and would be inhibited in audiovisual speech, where redundancy is present and would direct attention to amodal aspects of speech such as tempo, rhythm, affect, or prosody. Further, we expected that face discrimination would also be enhanced in an asynchronous audiovisual speech control condition. This condition provides the same amount (auditory plus visual) and types (auditory and visual) of stimulation as synchronous audiovisual speech; however, it provides no intersensory redundancy and thus should not direct attention away from the face to amodal properties such as tempo, rhythm, affect, or prosody. If the basis for attenuated face perception in synchronous audiovisual speech is in fact intersensory redundancy, then face discrimination in asynchronous audiovisual speech, where intersensory redundancy is absent, should be evident and should be equivalent to that in unimodal visual speech.

Thus, in Experiment 1, we examined whether infants 2 months of age could discriminate between the faces of two unfamiliar women in the context of intersensory redundancy (synchronous audiovisual speech) versus no intersensory redundancy (unimodal visual speech; asynchronous audiovisual speech). Further, we hypothesized that across development, infants' increased efficiency and flexibility of attention should lead to detection of modality-specific properties in both redundant audiovisual and nonredundant stimulation. Thus, face perception should be evident in older infants in the context of both unimodal visual and audiovisual synchronous speech. Since faces are such an important and prevalent type of stimulation, it may be that this improvement occurs rapidly across development. If so, differences should be observed even in cross-sectional data. In Experiment 2, we thus examined whether infants 3 months of age could discriminate between the faces of the same two unfamiliar women engaged in unimodal visual versus synchronous audiovisual speech.

We chose the ages of 2 and 3 months for the present studies on the basis of prior research assessing infant discrimination between two unfamiliar faces of women. Infants 4 months but not 2 months of age showed evidence of discriminating and matching the faces and voices of two unfamiliar women speaking (Bahrick et al., 2005), and thus we expected developmental change to occur between these ages. Further, 2-month-old infants were expected to show evidence of face discrimination in unimodal visual speech in the present study, given that infants of this age discriminated between two static live faces of unfamiliar women even with external cues masked (Blass & Camp, 2004) and between two dynamic silent faces of unfamiliar women in a control condition of the study described above (Bahrick et al., 2005). These findings also suggest that visual acuity is sufficiently developed by 2 months of age to support discrimination between the faces of unfamiliar women. Further, in Bahrick et al. (2005), 4-month-old infants discriminated unfamiliar faces in a task that was more demanding than the present task (requiring discrimination and

memory for two faces, two voices, and the relation between the two). Thus, for Experiment 2, we hypothesized that somewhat younger infants (i.e., 3-month-olds) might have had sufficient experience with faces to discriminate faces of unfamiliar women even in the context of intersensory redundancy from audiovisual speech.

Experiment 1: Discrimination of Faces in Synchronous Audiovisual, Asynchronous Audiovisual, and Unimodal Visual Speech at 2 Months of Age

In this study, 2-month-old infants were habituated with the face of a woman speaking under conditions of nonredundant speech (unimodal visual, silent speech or asynchronous audiovisual speech) versus redundant, audiovisual speech (synchronous face and voice). Consistent with predictions of the IRH, face discrimination, based primarily on modality-specific visual information, should be enhanced in unimodal visual speech, where no intersensory redundancy is available and attention is thus free to focus on information specific to visual stimulation, such as facial features and their configuration. In contrast, face discrimination should be attenuated in synchronous audiovisual speech, because salient intersensory redundancy should attract attention to amodal properties such as rhythm, tempo, and prosody at the expense of information about the appearance of the face. (Note that the terms *enhanced* and *impaired* or *attenuated* are used here as relative terms and refer to comparisons between infant detection of the same event properties across conditions that provide versus do not provide intersensory redundancy.) Several studies have already confirmed that amodal properties available in synchronous audiovisual speech and nonspeech events attract infant attention and are discriminated more readily than the same properties in nonredundant unimodal speech and nonspeech (Bahrick & Lickliter, 2000; Bahrick, Flom, & Lickliter, 2002; Castellanos et al., 2004; Flom & Bahrick, 2007; Flom, Gentile, & Pick, 2008). What is not known is whether intersensory redundancy provided by synchronous audiovisual speech will interfere with face discrimination in early development and, if so, whether this interference will be due to the intersensory redundancy per se. An alternative hypothesis also tested here is that synchronous audiovisual stimulation from speaking faces impairs young infants' face discrimination because the speech per se is interfering and/or the audiovisual speech provides a greater overall amount of stimulation than the unimodal visual speech, which interferes with attention to facial configuration. To distinguish between these hypotheses, an asynchronous control condition was included, which presented the same speaking faces and soundtracks as the synchronous speech condition, however the movements of the face were out of synchrony with the audible speech. This provided the same overall amount and type of stimulation as the synchronous condition but eliminated intersensory redundancy. If interference from intersensory redundancy impairs face perception in audiovisual speech (relative to unimodal visual speech), then face discrimination should also be enhanced even in the presence of speech as long as it is not synchronized with the movements of the face (thus eliminating intersensory redundancy). In contrast, if speech or the greater amount of stimulation it provides interferes with face perception (independent of intersensory redundancy), then infants should show impaired face

perception in the presence of asynchronous audiovisual speech (relative to performance in the visual speech condition).

Method

Participants. Forty-eight 2-month-old infants (15 girls and 33 boys) with a mean age of 63.3 days ($SD = 3.8$) participated. Infants in this and the subsequent studies reported here were all healthy, had a gestation period of at least 39 weeks, and were primarily from middle-class families. Forty were Hispanic, 7 were Caucasian not of Hispanic ethnicity, and 1 was African American. Twenty-two (46%) of the infants heard English or English and Spanish spoken in the home, 18 (38%) heard primarily Spanish, 3 (6%) heard another language, and the language of 5 (10%) was unknown/not reported by the parent. Forty additional infants participated, but their data were excluded from the analyses due to fussiness ($n = 2$ in the synchronous audiovisual condition and $n = 5$ in the unimodal visual condition), falling asleep ($n = 3$ in the synchronous audiovisual condition), experimenter error ($n = 1$ in the synchronous audiovisual condition, $n = 1$ in the unimodal visual condition), equipment failure ($n = 1$ in the synchronous audiovisual condition and $n = 1$ in the unimodal visual condition), external interference ($n = 1$ in the unimodal visual condition), and failure to meet the fatigue ($n = 4$ in the synchronous audiovisual condition, $n = 1$ in the unimodal visual condition, and $n = 3$ in the asynchronous audiovisual condition) or habituation criteria ($n = 8$ in the synchronous audiovisual condition, $n = 3$ in the unimodal visual condition, and $n = 3$ in the asynchronous audiovisual condition). Analyses indicated that there were no significant differences across the three conditions in the number of infants whose data were rejected, $\chi^2(2) = 1.88, p = .39$.

Stimulus events. Dynamic color video displays of two women reciting the nursery rhymes “Mary Had a Little Lamb,” and “Jack and Jill,” in a continuous loop in infant-directed speech served as our stimulus events. The video displays depicted the actresses’ faces and shoulder areas (see Figure 1). Both women were Caucasian, and given the large number of Hispanic participants in our area, one woman was of Hispanic ethnicity and the other was non-Hispanic. One woman had fair skin and straight light brown hair. The other woman had olive skin and curly dark brown hair. Three 30-s films were created for each woman, one for

each condition (synchronous audiovisual, unimodal visual, asynchronous audiovisual). The visual events were identical across conditions, and only their relation to the soundtrack differed. The stimulus events for the synchronous audiovisual habituation and test trials depicted a moving face producing natural and synchronous infant-directed speech. The stimulus events for the unimodal visual habituation and test trials depicted the same visual event but with the soundtrack removed (the moving face speaking silently). The stimulus events for the asynchronous audiovisual habituation and test trials depicted the moving faces; however, the soundtrack was out of synchrony with the movements of the faces. We accomplished this by delaying the soundtrack from the 30-s nursery rhyme sequence by 15 s with respect to the visual event. A control event was also used depicting a plastic green and white toy turtle whose arms spun while making a whirring sound.

Apparatus. Infants sat in a standard infant seat facing a color television monitor (Sony KV-20520) approximately 55 cm away. Black curtains surrounded the television monitor to obscure extraneous stimuli, and two 1.5-cm apertures allowed trained observers to view the infants’ visual fixations. Four Panasonic video desks (DS545 and AG7750) were used to play the stimulus events. Observers, unaware of the infants’ condition, depressed buttons on a joystick recording the length of the infants’ visual fixations. The joystick was connected to a computer that collected the data online. The observations of the primary observer controlled the stimulus presentations, and those of the secondary observer were recorded for use in calculating interobserver reliability.

Procedure. Infants were tested in an infant-controlled habituation procedure (similar to procedures used by Bahrick & Lickliter, 2000, 2004) to determine whether they could detect a change in the face (from familiar to novel) under conditions of synchronous audiovisual ($n = 16$), unimodal visual ($n = 16$), or asynchronous audiovisual ($n = 16$) exposure. Infants in the synchronous audiovisual condition received an audible and visible presentation of the woman speaking in synchrony with the movements of her face throughout the procedure. Infants in the asynchronous audiovisual condition received the same audible and visible presentation, but the woman’s face and voice were out of synchrony throughout the procedure. Infants in the unimodal visual condition received the same visual presentation of the woman’s face but with no audible soundtrack. The faces were counter-

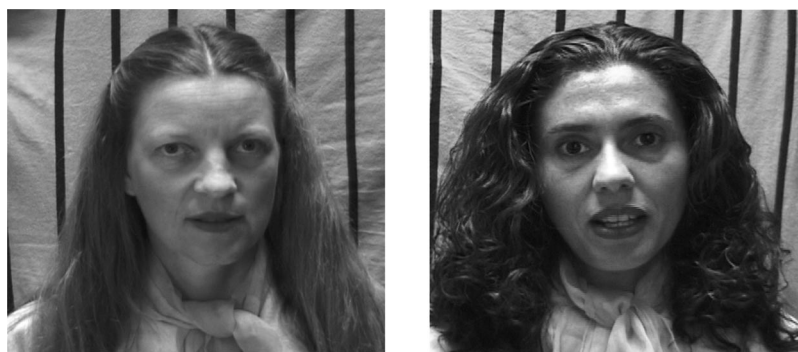


Figure 1. Static images of the two dynamic face events used in Experiments 1 and 2. Infants were habituated to one face and tested with the other face (order counterbalanced across infants) in sequential habituation and test trials.

balanced across infants so that half of the infants in each condition were habituated to the face of one woman and half were habituated to the face of the other woman.

In the infant-controlled habituation procedure, the face of one woman was presented throughout habituation and the face of the other woman was presented during test trials. The habituation procedure began with a control trial depicting a toy turtle and continued with four mandatory habituation trials. All trials began when infants visually fixated on the image and terminated when infants looked away for 1 s or after 45 s had elapsed. Additional trials of the same event were presented (up to a total of 20 habituation trials) until the infant's visual fixation level decreased and reached the habituation criterion, a decrement of 50% or greater on two consecutive trials relative to the infant's fixation level on the first two trials of the habituation sequence. Once the habituation criterion was met, two no-change posthabituation trials, identical to the habituation trials, were presented. Infants, under their respective conditions, then received two sequential test trials depicting a novel woman speaking. The nursery rhymes were spoken in a continuous loop, and thus the point in the rhyme where each new trial began was determined by where the previous trial ended. Infants in the synchronous audiovisual condition received two test trials depicting a novel face speaking in temporal synchrony with the familiar voice, those in the asynchronous audiovisual condition received two test trials depicting a novel face speaking out of synchrony with the familiar voice, and those in the unimodal visual condition received two test trials depicting a novel face speaking silently. Thus, the only change from habituation to test in each condition was the presentation of a novel face. Discrimination of the novel face was inferred when infants' visual fixation during the test trials depicting the novel face showed a significant increase relative to their visual fixation during the no-change post habituation trials depicting the familiar face (visual recovery). A final control trial depicting the toy turtle concluded the testing session.

To examine if infants had become fatigued, we compared their visual fixations to the initial and final control trials. Infants were

judged to be fatigued if their visual fixation to the final control trial was less than 20% of their fixation level to the initial control trial (see Bahrick et al., 2006, 2010). Two observers monitored 25 (52%) of the infants. The Pearson product-moment correlation between the visual fixation scores of the two observers averaged .94 ($SD = .12$) and served as our measure of interobserver reliability.

Results

Figure 2 depicts the mean visual recovery to the novel face for infants in the synchronous audiovisual, unimodal visual, and asynchronous audiovisual speech conditions. To address the primary research question, whether infants in each condition could discriminate between the novel and familiar faces, we conducted single-sample t tests on the mean visual recovery scores against the chance value of zero. Results were consistent with our predictions. Infants in the conditions that provided no intersensory redundancy (unimodal visual and asynchronous audiovisual) discriminated the novel from the familiar face, $t(15) = 4.37, p = .001$, and $t(15) = 5.26, p = .0001$, respectively, but those in the condition that provided intersensory redundancy (synchronous audiovisual) did not, $t(15) = 1.67, p > .10$.

To compare face discrimination across conditions, we conducted a one-way analysis of variance (ANOVA) on visual recovery with condition (synchronous audiovisual, unimodal visual, asynchronous audiovisual) and stimulus event (Woman A, Woman B) as the between-subjects factors. Results revealed a main effect of condition, $F(2, 45) = 5.55, p = .007, \eta^2 = .11$. Planned pairwise comparisons indicated that infants in the unimodal visual condition demonstrated significantly greater visual recovery to the novel face than did infants in the synchronous audiovisual condition ($p = .05$). These findings support predictions of the IRH and demonstrate unimodal facilitation for face perception. Planned pairwise comparisons also indicated that infants in the asynchronous audiovisual condition showed significantly greater visual recovery to the novel face than did infants in the synchronous

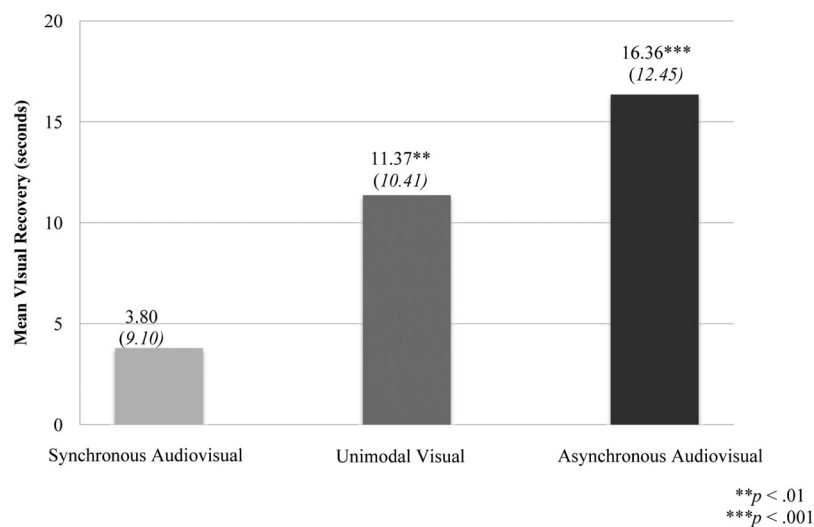


Figure 2. Two-month-old infants ($N = 48$): Mean visual recovery (and SD) to a novel face as a function of stimulus condition (synchronous audiovisual, unimodal visual, asynchronous audiovisual).

audiovisual condition ($p = .002$). Given that these two conditions provide the same amount and type of stimulation and only the face–voice synchrony differed, these findings highlight the interfering role of intersensory redundancy (synchrony) in face discrimination at this age. No main effects of stimulus event (Woman A, Woman B) or interaction between stimulus event and condition (synchronous audiovisual, unimodal visual) were found ($ps > .10$), indicating that visual recovery did not differ as a function of which woman infants viewed during habituation. For all subsequent analyses, we collapsed across stimulus events.

Further, the primary language spoken in the home (English or English/Spanish vs. other) was not an important factor affecting whether or not speech interfered with face discrimination. Face discrimination was enhanced in the presence of English speech when it was asynchronous and impaired when it was synchronous despite the fact that some infants heard English and others heard Spanish or another language in the home ($ps > .1$). This is not surprising given that detection of temporal synchrony is known to be a basic perceptual skill underlying audiovisual speech perception and that infants 6 months of age and younger have been found to have broadly tuned intersensory processing skills and can match nonnative faces and vocalizations even at birth (Lewkowicz, 2010; Lewkowicz & Ghazanfar, 2009; Lewkowicz, Leo, & Simion, 2010).

Secondary analyses were conducted to compare infants' performance on four habituation measures (mean baseline looking, mean number of habituation trials, mean posthabituation looking, mean total processing time) across conditions. Table 1 presents infant performance on each of these measures along with their test trial looking time. One-way ANOVAs were conducted on each of the four habituation measures with condition (synchronous audiovisual, unimodal visual, asynchronous audiovisual) as the between-subjects factor. Results revealed no significant main effects of condition for any of the habituation measures (all $ps > .05$); however, the main effect for processing time was marginally significant, $F(2, 45) = 3.01, p = .06, \eta^2 = .02$. Pairwise comparisons indicated significantly greater processing time for infants in the asynchronous audiovisual condition than for infants in the unimodal visual condition ($p = .02$) and no other differences (all $ps > .05$). However, given that visual recovery did not differ between these two groups despite their differences in processing time, and that both showed evidence of face discrimination, it is

clear that processing time differences between the unimodal visual and asynchronous audiovisual groups did not affect the main results.

These results indicate that face discrimination is facilitated in conditions where intersensory redundancy is absent (unimodal visual and asynchronous audiovisual presentations) and impaired in the context of intersensory redundancy from synchronous audiovisual speech, and that these effects were independent of differential exposure to the faces across conditions. Infants in both conditions where intersensory redundancy was eliminated displayed robust evidence of face discrimination. Moreover, this facilitation of face discrimination was evident even in the unimodal visual condition despite the fact that infants in this condition showed the least amount of processing time.

Together, the findings indicate that by 2 months of age, face perception is impaired in synchronous audiovisual speech because intersensory redundancy interferes with attention to the visual appearance of the face. In contrast, face perception is enhanced in unimodal visual stimulation where intersensory redundancy is absent and thus cannot compete for attention. In the asynchronous audiovisual control condition, in which the amount and type of stimulation were equal to those of the synchronous audiovisual condition, there was no evidence of interference from the presence of the voice or the overall amount of stimulation provided by the combination of auditory and visual speech. This eliminated these alternative explanations for impaired face perception in synchronous audiovisual speech. Together these results demonstrate the central role of intersensory redundancy in attention and perception of faces in early development.

Experiment 2: Discrimination of Faces in Synchronous Audiovisual Versus Unimodal Visual Speech at 3 Months of Age

In Experiment 1, 2-month-old infants showed no evidence of face discrimination during synchronous audiovisual speech, where intersensory redundancy is present, but showed significant face discrimination in both unimodal visual and asynchronous audiovisual speech, conditions where intersensory redundancy is not available. According to the IRH, across development, infants' increased experience with multimodal events along with improvements in their efficiency of attention and perceptual flexibility lead

Table 1
Means (and Standard Deviations) for Visual Fixation (in Seconds) for Baseline, Trials to Habituation, Posthabituation, Processing Time, and Test Trials as a Function of Age and Condition

Condition	Baseline	Trials to habituation	Posthabituation	Processing time	Test
Experiment 1 (age = 2 months)					
Synchronous audiovisual	31.58 (11.75)	9.19 (4.52)	9.62 (9.52)	217.71 (82.64)	13.42 (10.45)
Unimodal visual	29.25 (10.37)	10.75 (4.09)	7.94 (4.64)	196.30 (83.92)	19.31 (12.40)
Asynchronous audiovisual	30.55 (10.78)	11.19 (3.35)	5.91 (3.42)	285.17 (142.85)	22.27 (13.12)
Experiment 2 (age = 3 months)					
Synchronous audiovisual	46.93 (12.64)	8.06 (2.79)	10.67 (5.15)	298.84 (185.47)	21.90 (14.44)
Unimodal visual	45.52 (15.32)	9.00 (3.31)	8.91 (6.56)	301.30 (166.63)	23.18 (15.57)

Note. Baseline = first two habituation trials; Trials to habituation = number of habituation trials to reach criterion; Posthabituation = two no-change trials following habituation, reflecting final interest level; Processing time = total number of seconds fixating the habituation events.

to detection of both amodal and modality-specific properties in redundant audiovisual and nonredundant (unimodal visual, asynchronous audiovisual) stimulation (Bahrick, 2010; Bahrick & Lickliter, 2012; Bahrick et al., 2010). The present study tested this developmental prediction by assessing face discrimination in older infants under conditions of unimodal visual versus synchronous audiovisual stimulation, identical to Experiment 1. In addition to the developmental improvements observed in face perception between 2 and 4 months of age (Bahrick et al., 2005), 3-month-olds were selected because they have a great deal of experience with faces and synchronous audiovisual speech by this age and we expected that development would progress rapidly in this domain. We thus predicted that by 3 months of age, infants would show evidence of face discrimination even in the context of intersensory redundancy from audiovisual speech.

Method

Participants. Thirty-two 3-month-old infants (17 girls and 15 boys) with a mean age of 92.69 days ($SD = 5.52$) participated. Twenty-nine were Hispanic and 3 were Caucasian, not of Hispanic ethnicity. Nine (28%) of the infants heard English or English and Spanish spoken in their homes, 16 (50%) of the infants heard primarily Spanish, and for 7 (22%) the language spoken in the home was unknown or not reported by the parent. Fifteen additional infants participated, but their data were excluded from the analyses because of fussiness ($n = 1$ in the synchronous audiovisual condition and $n = 3$ in the unimodal visual condition), equipment failure ($n = 1$ in the unimodal visual condition), and failure to meet the fatigue ($n = 4$ in the synchronous audiovisual condition and $n = 3$ in the unimodal visual condition) or habituation criteria ($n = 1$ in the synchronous audiovisual condition and $n = 2$ in the unimodal visual condition).

Stimulus events, apparatus, and procedure. The stimulus events, apparatus, and procedures were identical to those of Experiment 1. Two observers monitored 11 (34%) of the infants, and a Pearson product-moment correlation between the visual fixation scores of the two observers averaged .95 ($SD = .05$).

Results

Figure 3 depicts the mean visual recovery to the novel face as a function of stimulus condition (synchronous audiovisual, unimodal visual) for 3-month-old infants. Single-sample t tests were conducted against the chance value of zero to determine whether infants showed significant visual recovery to the novel face. Results confirmed our predictions and indicated that 3-month-old infants discriminated the novel from the familiar face following both unimodal visual, $t(15) = 3.66$, $p = .002$, and synchronous audiovisual, $t(15) = 3.31$, $p = .005$, habituation.

A one-way ANOVA was conducted on visual recovery scores with stimulus condition (unimodal visual, synchronous audiovisual) as the between-subjects factor to compare discrimination across conditions. As predicted, results revealed no main effect of condition, $F(1, 30) = 0.36$, $p = .56$, $\eta^2 = .01$, indicating that 3-month-olds discriminated the novel from the familiar face across both conditions. Further, secondary analyses indicated no difference in face discrimination as a function of language spoken at home (English or English/Spanish vs. other; $ps > .1$) and no differences between conditions for any of the habituation measures (see Table 1). These findings indicate that by 3 months of age, infants have had sufficient experience with faces and individuals speaking that they can discriminate between faces even in the context of salient intersensory redundancy provided by audiovisual speech.

General Discussion

This research explored the role of intersensory redundancy generated by naturalistic audiovisual speech in the development of infant face discrimination. According to the IRH, intersensory redundancy available in synchronous audiovisual speech should impair face discrimination by focusing attention on highly salient amodal properties at the expense of modality-specific properties that underlie face discrimination. In contrast, when intersensory redundancy is eliminated by presenting unimodal visual speech, attention should be free to focus on facial features and their

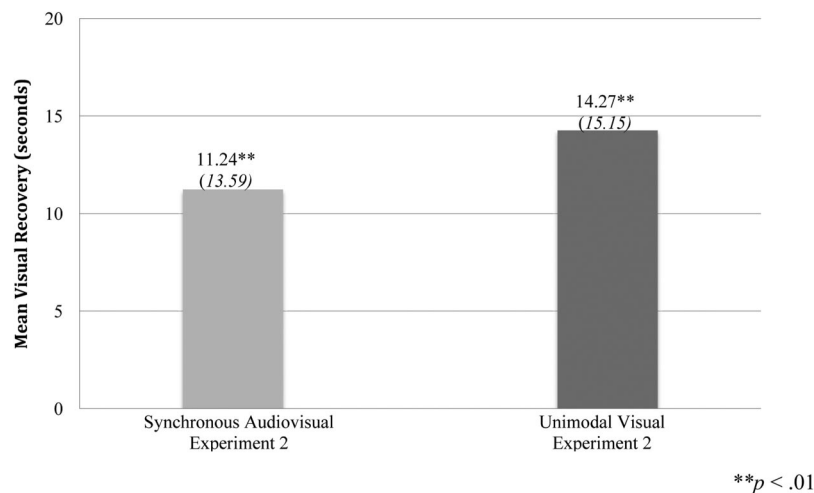


Figure 3. Three-month-old infants ($N = 32$): Mean visual recovery (and SD) to a novel face as a function of stimulus condition (synchronous audiovisual, unimodal visual).

configuration, promoting face discrimination. Findings were consistent with these predictions and demonstrated that at the age of 2 months (Experiment 1) face perception was impaired in the context of audiovisual speech, where face–voice synchrony provides intersensory redundancy, and was enhanced in unimodal visual speech, where no intersensory redundancy is available to compete for attention. Following habituation to a woman speaking in synchrony with her voice, 2-month-old infants showed no visual recovery to the face of a novel woman speaking. In contrast, when intersensory redundancy was eliminated by showing the same face speaking silently, 2-month-olds showed significant visual recovery to the face of a novel woman speaking silently. Results of the asynchronous control condition confirmed that the basis for this effect was attentional competition from highly salient intersensory redundancy rather than differential amounts or types (visual only vs. visual plus auditory) of stimulation. When intersensory redundancy was eliminated while controlling for the amount and type of stimulation by presenting asynchronous faces and voices, 2-month-olds also showed significant evidence of face discrimination. Visual recovery to a novel face was evident in the context of asynchronous, but not synchronous, audiovisual speech. Thus, the voice itself was not distracting, and infants' failure to discriminate was not due to an inability to discriminate faces in the context of voices; rather, the synchronous temporal relation of the voice with the movements of the face was the basis for interference.

Further, our main pattern of findings across conditions was unrelated to the amount of time infants spent processing the events during habituation. Analyses indicated a marginally significant main effect of processing time across conditions; however, only the asynchronous audiovisual and unimodal visual conditions differed, with greater processing time in the asynchronous than in the unimodal visual condition. Despite their different processing times, infants in the two conditions did not differ from one another in their visual recovery scores, and infants in both conditions showed significant visual recovery to the novel face. Thus, longer looking to the faces did not result in greater face discrimination, and even unimodal visual exposure (associated with the least processing time) resulted in face discrimination. This further supports the conclusion that face discrimination is a function of attentional allocation to different properties of stimulation: Intersensory redundancy interferes with face discrimination by focusing attention on amodal properties, and unimodal visual exposure enhances face discrimination because, in the absence of competition from salient intersensory redundancy, attention is free to focus on visual features and their configuration.

Experiment 2 revealed that development in the domain of face perception occurs rapidly. At the age of 2 months, infants discriminated between two faces of unfamiliar women only under conditions that provided no intersensory redundancy (unimodal visual and asynchronous audiovisual speech). In contrast, by the age of 3 months, infants were able to discriminate the faces even in the context of interference from highly salient intersensory redundancy and showed significant visual recovery to the novel face under both the unimodal visual and synchronous audiovisual speech conditions. This suggests that the additional month of experience with faces and voices enhanced face processing skills, effectively reducing the difficulty of the discrimination task and promoting more flexible attentional allocation. These findings converge with those of our prior study demonstrating developmen-

tal change between the ages of 2 and 4 months in infant perception of unfamiliar faces of women (Bahrick et al, 2005). Together, these findings suggest that as infants gain experience with faces and voices of speaking people, their attention becomes more flexible and extends to modality-specific properties that support face discrimination, even in the context of intersensory redundancy from audiovisual speech. It is important to note, however, that although this developmental sequence (modality-specific properties detected in nonredundant stimulation in early development and extending to redundant stimulation in later development) is predicted to be invariant across domains, the age at which this transition takes place would depend on the difficulty of the discrimination task in relation to the skills of the perceiver (see Bahrick & Lickliter, 2012; Bahrick et al., 2010). Just as older infants reverted to patterns of intersensory facilitation characteristic of younger infants when the difficulty of a tempo discrimination task was increased (Bahrick et al., 2010), older infants should revert to patterns of intersensory interference characteristic of younger infants if task difficulty is increased in the domain of face perception. Thus, even older infants and children should show evidence of unimodal facilitation followed by a transition to more flexible face processing for tasks that are challenging.

These findings are the first to demonstrate the significant role of intersensory redundancy in infant face perception. Specifically, they demonstrate that intersensory redundancy generated by naturalistic speech interferes with discrimination of faces in early development. Intersensory redundancy is highly salient and promotes attention to amodal properties of stimulation such as synchrony, rhythm, tempo, affect, and prosody of speech (Bahrick, 2010; Bahrick & Lickliter, 2003, 2012; Castellanos, 2007; Castellanos & Bahrick, 2008; Castellanos et al., 2004; Flom & Bahrick, 2007). This gives rise to attentional salience hierarchies such that infants show earlier, deeper, and/or more complete processing of amodal properties than other properties in multimodal stimulation (see Reynolds, Bahrick, Lickliter, & Guy, *in press*, for event-related potential evidence of greater depth of processing for redundant than nonredundant audiovisual speech events). Across exploratory time, attention progresses along the salience hierarchy such that more salient aspects of stimulation are processed first. Once the most salient properties are processed, attention is then allocated to less salient properties of stimulation, including modality-specific properties that support face discrimination in audiovisual speech (see Bahrick & Newell, 2008). Across development, attention becomes more efficient and can progress down the salience hierarchy more quickly, allowing older infants to process both the more and less salient properties in a single episode of exploration. In contrast, when no intersensory redundancy is available to compete for attention, such as in unimodal visual face displays or asynchronous faces and voices, attention is free to focus on visual information supporting face discrimination. Thus, in early development, when attentional resources are most limited, intersensory redundancy available in audiovisual speech interferes with discrimination of unfamiliar faces (when the task is relatively difficult in relation to the skills of the perceiver). It is not yet known to what extent experience with particular faces can speed up this developmental process and the progression of infant attention along the salience hierarchy, promoting discrimination of faces in both unimodal and multimodal stimulation by the age of 2 months or earlier.

The present findings complement those of our prior studies that demonstrated another important factor, the role of motion and repetitive action in shaping infant attentional salience hierarchies and in turn affecting face perception (Bahrick, Gogate, & Ruiz, 2002; Bahrick & Newell, 2008). By 5 months, everyday activities (e.g., brushing hair, brushing teeth, or blowing bubbles) are highly salient and processed at the expense of the faces engaged in the activity. Whether more subtle movements, such as those made during speech, would interfere with discrimination of faces in very young infants when no intersensory redundancy was present (e.g., silent speech) is not known. It is possible that prior to 2 months, when attentional resources are even more limited, even the movements of speech could be sufficient to capture attention and promote processing at the expense of the faces engaged in action. Together with the present findings, these results suggest that infant face perception skills have been overestimated in the literature, which has primarily focused on perception of static and/or silent faces (see Bahrick, Gogate, & Ruiz, 2002; Bahrick & Newell, 2008; Otsuka et al., 2009, for discussion). These findings likely reflect discrimination of silent, static faces but may not generalize well to face perception in the dynamic, multimodal environment. Our results suggest that infant attention to faces in the natural environment is highly variable, depends on context and task difficulty in relation to the skills of the perceiver, and is negatively affected by the presence of other stimulus properties that compete for attention, including intersensory redundancy and repetitive actions. Thus, in the natural environment, the focus of attention may change dynamically in real time as the conditions of stimulation shift. While people are speaking, infants' attention may be more focused on amodal properties of speech; when people are silent but gesturing, infants' attention would likely shift to the nature of movement at the expense of the appearance of individuals; and when people are silent with minimal motion, infants' attention may shift to modality-specific visual properties including the facial configuration. The implications of these findings for optimizing face perception are clear. In early infancy, face perception and discrimination should be optimal when faces are experienced in unimodal visual stimulation, without interference from synchronous audiovisual speech or highly distinctive or salient motions. For example, motions such as gestures, which are typically synchronized with speech, would promote attention to amodal properties and would likely interfere with face discrimination. More minimal motion (such as gradual changes in facial expression or head movement), however, appears to enhance face discrimination over static presentations, at least for unfamiliar faces of women by infants 3 to 4 months of age, as it provides invariant information for facial features across changes in perspective over time (Otsuka et al., 2009).

Moreover, these findings demonstrate that face perception is governed by the same perceptual principles that govern perception of other objects and events (see also Gauthier & Nelson, 2001; Nelson, 2003). Consistent with tenets of the IRH, face perception, which relies on information specific to vision, shows evidence of unimodal facilitation, similar to other events that rely on information specific to vision (e.g., orientation of hammers tapping; Bahrick et al., 2006; Flom & Bahrick, 2010). Unimodal facilitation of face perception can be seen as an example of a general principle of early development applicable across species, from human to avian infants (Vaillant et al., 2012). Rather than comprising a special

class of stimuli that are salient and preferred independent of experience (Goren et al., 1975; Johnson & Morton, 1991), facial features and their arrangement appear to be salient in early development (e.g., prior to 3 months) only in the absence of attentional competition from intersensory redundancy or repetitive actions. However, the transition from discriminating faces only in the context of no intersensory redundancy (e.g., silent visual) to more flexible discrimination of faces in contexts of both redundancy and no redundancy (e.g., synchronous audiovisual speech and silent visual presentations) occurs early (between the ages of 2 and 3 months for the face events used in the present study) compared with discrimination of object events tested thus far (e.g., between 5 months and 8 or 9 months, Bahrick et al., 2006; Flom & Bahrick, 2010). This is likely due to infants' high degree of familiarity and variability of exposure with faces. Together, these findings add to a growing body of research demonstrating that the development of face perception is organized by general principles of perceptual development, including the operation of attentional salience hierarchies, perceptual narrowing, and interference from intersensory redundancy and repetitive motions (Bahrick, Gogate, & Ruiz, 2002; Bahrick & Newell, 2008; Gauthier & Tarr, 1997; Lewkowicz & Ghazanfar, 2006; Pascalis, de Haan, & Nelson, 2002).

It is not yet clear to what extent intersensory redundancy interferes with face perception in later development. Attentional salience hierarchies have the greatest effect on attention and perception when attentional resources are limited, tasks are difficult in relation to the skills of the perceiver, or memory load is high (Adler et al., 1998; Bahrick, 2010; Bahrick et al., 2010; Craik & Byrd, 1982; Craik et al., 2010). Our prior research demonstrates that even when older infants can detect amodal properties in both unimodal visual and synchronous audiovisual stimulation, intersensory facilitation can be reinstated if the task is made more difficult (Bahrick et al., 2010). Thus, the present findings are likely to have more broad implications for face perception across the life span. When face discrimination is difficult, or under conditions of high stress, such as witnessing a crime, face discrimination should also be impaired in the presence of intersensory redundancy from audiovisual speech. In contrast, for simpler face discrimination tasks, such as detecting large differences between individuals (e.g., on the basis of gender or age), or for discriminating among familiar faces, face discrimination may be evident in very young infants, even in the context of interference from intersensory redundancy. Further research with older children (Bahrick, Krogh-Jespersen, Argumosa, & Lopez, 2012) and adults is currently under way in our lab to reveal more about the conditions that impair versus enhance face perception across the life span.

References

- Adler, S. A., Gerhardstein, P., & Rovee-Collier, C. (1998). Levels-of-processing effects in infant memory? *Child Development, 69*, 280–294. doi:10.2307/1132164
- Bahrick, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development, 59*, 197–209. doi:10.2307/1130402
- Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology, 79*, 253–270. doi:10.1006/jecp.2000.2588

- Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In G. Bremner & T. D. Wachs (Eds.), *Blackwell handbook of infant development*. (2nd ed., pp. 120–166) Oxford, England: Blackwell Publishing. doi:10.1002/9781444327564.ch4
- Bahrick, L. E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology*, *41*, 352–363.
- Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development*, *73*, 1629–1643. doi:10.1111/1467-8624.00495
- Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face–voice relations. *Developmental Psychology*, *41*, 541–552. doi:10.1037/0012-1649.41.3.541
- Bahrick, L. E., Krogh-Jespersen, S., Argumosa, M. A., & Lopez, H. (2012). *Intersensory redundancy hinders face discrimination in preschool children: Evidence for visual facilitation*. Manuscript submitted for publication.
- Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*, 190–201. doi:10.1037/0012-1649.36.2.190
- Bahrick, L. E., & Lickliter, R. (2003). Intersensory redundancy guides early perceptual and cognitive development. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 153–187). New York, NY: Academic Press. doi:10.1016/S0065-2407(02)80041-6
- Bahrick, 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. doi:10.3758/CABN.4.2.137
- Bahrick, 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). Oxford, England: Oxford University Press. doi:10.1093/acprof:oso/9780199586059.003.0008
- Bahrick, 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. doi:10.1111/j.1467-7687.2009.00928.x
- Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, *13*, 99–102. doi:10.1111/j.0963-7214.2004.00283.x
- Bahrick, L. E., Lickliter, R., & Flom, R. (2006). Up versus down: The role of intersensory redundancy in development of infants' sensitivity to the orientation of moving objects. *Infancy*, *9*, 73–96. doi:10.1207/s15327078in0901_4
- Bahrick, L. E., Lickliter, R., Shuman, M. A., Batista, L. C., & Grandez, C. (2003, April). *Infant discrimination of voices: Predictions from the intersensory redundancy hypothesis*. Poster presented at the meeting of the Society for Research in Child Development, Tampa, FL.
- Bahrick, L. E., Moss, L., & Fadil, C. (1996). Development of visual self-recognition in infancy. *Ecological Psychology*, *8*, 189–208. doi:10.1207/s15326969eco0803_1
- Bahrick, L. E., & Newell, L. C. (2008). Infant discrimination of faces in naturalistic events: Actions are more salient than face. *Developmental Psychology*, *44*, 983–996. doi:10.1037/0012-1649.44.4.983
- Bartrip, J., Morton, J., & de Schonen, S. (2001). Responses to mother's face in 3-week to 5-month-old infants. *British Journal of Developmental Psychology*, *19*, 219–232. doi:10.1348/026151001166047
- Blass, E. M., & Camp, C. A. (2004). The ontogeny of face identity: I. Eight- to 21-week-old infants use internal and external face features in identity. *Cognition*, *92*, 305–327. doi:10.1016/j.cognition.2003.10.004
- Bushnell, I. W. R. (2001). Mother's face recognition in newborn infants: Learning and memory. *Infant and Child Development*, *10*, 67–74. doi:10.1002/icd.248
- Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*, *15*, 379–383. doi:10.1111/j.0956-7976.2004.00688.x
- Castellanos, I. (2007). *Intersensory redundancy educates human infants' attention to the prosody of speech* (Unpublished master's thesis). Florida International University, Miami, FL.
- Castellanos, I., & Bahrick, L. E. (2008, November). *Educating infants' attention to the amodal properties of speech: The role of intersensory redundancy*. Poster presented at the meeting of the International Society for Developmental Psychobiology, Washington, DC.
- Castellanos, I., Shuman, M. A., & Bahrick, 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.
- Castellanos, I., Vaillant-Molina, M., Lickliter, R., & Bahrick, L. E. (2006, October). *Intersensory redundancy educates infants' attention to amodal information during early development*. Poster presented at the meeting of the International Society for Developmental Psychobiology, Atlanta, GA.
- Cohen, L. B., & Cashon, C. H. (2001). Do 7-month-old infants process independent features or facial configurations? *Infant and Child Development*, *10*, 83–92. doi:10.1002/icd.250
- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). New York, NY: Plenum Press.
- Craik, F. I. M., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. *Psychology and Aging*, *25*, 968–979. doi:10.1037/a0020276
- Farzin, F., Charles, E., & Rivera, S. M. (2009). Development of multimodal processing in infancy. *Infancy*, *14*, 563–578. doi:10.1080/15250000903144207
- Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother–stranger face discrimination by the newborn. *Infant Behavior & Development*, *7*, 19–25. doi:10.1016/S0163-6383(84)80019-3
- Flom, R., & Bahrick, 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. doi:10.1037/0012-1649.43.1.238
- Flom, R., & Bahrick, L. E. (2010). The effects of intersensory redundancy on attention and memory: Infants' long-term memory for orientation in audiovisual events. *Developmental Psychology*, *46*, 428–436. doi:10.1037/a0018410
- Flom, R., Gentile, D. A., & Pick, A. D. (2008). Infants' discrimination of happy and sad music. *Infant Behavior & Development*, *31*, 716–728. doi:10.1016/j.infbeh.2008.04.004
- Frank, M. C., Slemmer, J., Marcus, G., & Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental Science*, *12*, 504–509. doi:10.1111/j.1467-7687.2008.00794.x
- Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*, 219–224. doi:10.1016/S0959-4388(00)00200-2
- Gauthier, I., & Tarr, M. J. (1997). Becoming a “greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, *37*, 1673–1682. doi:10.1016/S0042-6989(96)00286-6
- Gogate, L. J., & Hollich, G. J. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review*, *117*, 496–516. doi:10.1037/a0019049
- Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word comprehension: An ecological-dynamic

- systems view. *Developmental Science*, 4, 1–18. doi:10.1111/1467-7687.00143
- Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56, 544–549.
- Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infant's use of synchronized visual information to separate streams of speech. *Child Development*, 76, 598–613. doi:10.1111/j.1467-8624.2005.00866.x
- 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. doi:10.1016/0010-0277(91)90045-6
- Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development*. Oxford, England: Blackwell.
- Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition*, 108, 210–221. doi:10.1016/j.cognition.2007.12.001
- Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, 104, 105–114. doi:10.1016/j.jecp.2009.01.006
- Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, 126, 281–308. doi:10.1037/0033-2909.126.2.281
- Lewkowicz, D. J. (2004). Perception of serial order in infants. *Developmental Science*, 7, 175–184. doi:10.1111/j.1467-7687.2004.00336.x
- Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology*, 46, 66–77. doi:10.1037/a0015579
- Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences, USA*, 103, 6771–6774. doi:10.1073/pnas.0602027103
- Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, 13, 470–478. doi:10.1016/j.tics.2009.08.004
- Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, 15, 46–60. doi:10.1111/j.1532-7078.2009.00005.x
- Lewkowicz, D. J., & Lickliter, R. (Eds.). (1994). *Development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum.
- Lickliter, R., & Bahrick, L. E. (2004). Perceptual development and the origins of multisensory responsiveness. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory integration* (pp. 643–654). Cambridge, MA: MIT Press.
- Lickliter, R., Bahrick, L. E., & Honeycutt, H. (2002). Intersensory redundancy facilitates prenatal perceptual learning in bobwhite quail embryos. *Developmental Psychology*, 38, 15–23. doi:10.1037/0012-1649.38.1.15
- Lickliter, R., Bahrick, L. E., & Honeycutt, H. (2004). Intersensory redundancy enhances memory in bobwhite quail embryos. *Infancy*, 5, 253–269. doi:10.1207/s15327078in0503_1
- Lickliter, R., Bahrick, L. E., & Markham, R. G. (2006). Intersensory redundancy educates selective attention in bobwhite quail embryos. *Developmental Science*, 9, 604–615. doi:10.1111/j.1467-7687.2006.00539.x
- Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10, 419–422. doi:10.1111/1467-9280.00179
- Nelson, C. A. (2003). The development of face recognition reflects an experience-expectant and activity-dependent process. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 79–97). Hauppauge, NY: Nova Science.
- Otsuka, Y., Konishi, Y., Kanazawa, S., Yamaguchi, M. K., Abdi, H., & O'Toole, A. J. (2009). Recognition of moving and static faces by young infants. *Child Development*, 80, 1259–1271. doi:10.1111/j.1467-8624.2009.01330.x
- Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, 296, 1321–1323. doi:10.1126/science.1070223
- Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior & Development*, 18, 79–85. doi:10.1016/0163-6383(95)90009-8
- Reynolds, G. D., Bahrick, L. E., Lickliter, R., & Guy, M. W. (in press). Neural correlates of intersensory processing in five-month-old infants. *Developmental Psychobiology*.
- Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: Behavioral and functional magnetic resonance imaging data. *Journal of Cognitive Neuroscience*, 19, 1435–1452. doi:10.1162/jocn.2007.19.9.1435
- 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. doi:10.1002/icd.376
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, 25, 583–592. doi:10.3758/BF03211301
- Thompson, L. A., & Massaro, D. W. (1989). Before you see it, you see its parts: Evidence for feature encoding and integration in preschool children and adults. *Cognitive Psychology*, 21, 334–362. doi:10.1016/0010-0285(89)90012-1
- Vaillant, J., Bahrick, L. E., & Lickliter, R. (2012). *Detection of modality-specific pitch information in bobwhite quail chicks: Unimodal auditory facilitation and intersensory interference*. Manuscript submitted for publication.
- Vaillant-Molina, M., & Bahrick, L. E. (2012). The role of intersensory redundancy in the emergence of social referencing in 5.5-month-old infants. *Developmental Psychology*, 48, 1–9. doi:10.1037/a0025263
- Vaillant-Molina, M., Gutierrez, M. E., & Bahrick, L. E. (2005, November). *Infant memory for modality-specific properties of contingent and non-contingent events: The role of intersensory redundancy in self perception*. Poster presented at the meeting of the International Society for Developmental Psychobiology, Washington, DC.
- Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121, 437–456. doi:10.1037/0033-2909.121.3.437
- Ward, T. B. (1989). Analytic and holistic modes of categorization in category learning. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 387–419). Hillsdale, NJ: Erlbaum.
- Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, 18, 580–593. doi:10.1162/jocn.2006.18.4.580

Received January 10, 2012

Revision received August 13, 2012

Accepted August 16, 2012 ■