Infants' Intermodal Perception of Canine (*Canis familairis*) Facial Expressions and Vocalizations

Ross Flom
Brigham Young University

Heather Whipple
University of Minnesota

Daniel Hyde
Harvard University

From birth, human infants are able to perceive a wide range of intersensory relationships. The current experiment examined whether infants between 6 months and 24 months old perceive the intermodal relationship between aggressive and nonaggressive canine vocalizations (i.e., barks) and appropriate canine facial expressions. Infants simultaneously viewed static aggressive and nonaggressive expressions of the same canine and heard an aggressive or nonaggressive bark. Results indicate that 6-month-olds perceived the intermodal relationship for aggressive and nonaggressive barks and the appropriate expression. Results also revealed that in older but not younger infants, the initial or first looks were directed toward the appropriate expression and that older infants also looked proportionately longer to the incongruent expression during the latter half of the test trials. Findings are discussed in terms of perceptual narrowing and the effects of familiarity and experience.

Keywords: intersensory perception, perceptual narrowing, face perception, infancy

At birth human infants enter the world ready for social interaction and communication with others. Moreover, it is these social and communicative interactions that guide, and to some extent constrain, early perceptual, cognitive, and linguistic development.

It is well known that to even newborn infants, faces are salient, preferred over other stimuli (Barrera & Mauer, 1981; Easterbrook, Kisilevsky, Hains & Muir, 1999) and are readily discriminated (Fagan, 1972; Kleiner, 1987; Mauer & Young, 1983; Pascalis & de Schonen, 1994). Recently it has also been shown that over the course of development, infants' proclivity for face discrimination is influenced by their social environment. Pascalis, de Hann,, and Nelson (2002), for example, found that 6-month-olds are able to discriminate various monkey faces as well as human faces, whereas 9-month-olds can only discriminate different human faces. It is important to note, however, that if 6-month-olds receive 1–2 min per day of familiarization with photographs of monkey faces 3 months later, the now 9-month-olds can still discriminate unfamiliar and familiar monkey faces (Pascalis et al., 2005). Thus, including monkey faces in the infants' social environment allows infants to discriminate monkey faces at an older age than if they had not had experience with monkey faces. More broadly, this pattern of results suggests that infants' ability to discriminate faces mirrors their social and perceptual experiences (Nelson, 2001; Quinn, Yarr, Kuhn, Slater, & Pascalis, 2002).

In addition to their perception of faces, infants from birth are excellent perceivers of the human voice. Human fetuses, for example, attend to sounds and voices toward the end of gestation (Birnholtz & Benacerraf, 1983; Querleu, Renard, Boutteville, & Crepin, 1989), and infants show prenatal auditory learning of the maternal voice and prosody of speech (DeCasper & Fifer, 1980; DeCasper & Spence, 1986). Like face perception, infants' discrimination of speech also varies as a function of experience and exposure. Specifically, 4- to 6-month-olds discriminate speech sounds in their native and nonnative languages; however, by 10–12 months of age, infants can only discriminate speech sounds common to their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). More recently, however, Kuhl, Tsao, and Liu (2003) have shown that providing 9-month-olds with exposure to nonnative phonemes extends their discrimination of these phonemes to 12 months of age.

In sum, these results show that in the first months of life, infants discriminate faces and speech sounds under a diverse range of conditions. These results also suggest that infants' capacity to discriminate faces and speech sounds changes over the course of development: Younger, but not older, infants discriminate a wider range of speech sounds and faces. Finally, if infants do not continue to receive exposure to a particular language or the faces of a given species, infants may lose the ability to discriminate those speech sounds or faces.

Infants' discrimination of faces and the sounds of speech are certainly important in terms of their perceptual and cognitive development. Just as important, however, is the ability to detect which sights and sounds go together. Most objects and events, including other people, can be explored or experienced with more than one sense modality (i.e., they are multimodal), and the ability to perceive the relation between different types of sensory information is necessary for infants to arrive at a veridical perception and subsequent understanding of their world. Fortunately, by the age of 2 months, infants detect voice–lip synchrony (Dodd, 1979), and by 4 months, infants match audible speech and visible lip movements on the basis of spectral information in vowel sounds (Kuhl & Meltzoff, 1984; Patterson & Werker, 1999). Between the ages of 5 and 7 months, infants match faces and voices on the basis of affect (Soken & Pick, 1992; Walker 1982; Walker-Andrews, 1997), gender (Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991), and age of speaker (Bahrick, Netto, & Hernandez-Reif, 1998). Recent work by Lewkowicz and Ghazanfar (2006) showed that younger but not older infants are able to make intersensory matches for unfamiliar macaque faces and calls. Specifically, they found that 4- and 6-month-olds, but not 8- and 10-month-olds, match a macaque threat or coo call to the appropriate facial expression (Lewkowicz & Ghazanfar, 2006).

The fact that Lewkowicz and Ghazanfar (2006) found a decline of intersensory perception is important as it parallels a similar pattern of results in early visual and auditory development. In addition, the results of Lewkowicz and Ghazanfar (2006) are groundbreaking as they are the first to show cross-species intersensory matching by human infants. Unfortunately, the results of Lewkowicz and Ghazanfar (2006) do not reveal what information infants are using in perceiving the intersensory relation. One possibility is infants are using the temporal synchrony between a given call with the mouth closing. In Lewkowicz and Ghazanfar's study, the *onset* of the call was temporally synchronous with the mouth opening of the matching and nonmatching face, whereas the *offset* of the call was synchronous only with the matching face. Another possibility—and the possibility explored in the current study—is that infants perceive what is communicatively or affectively common across the auditory and visual information. Just as infants around 6 months of age match faces and voices on the basis of affect, gender, and age of the speaker (Bahrick, Netto, & Hernandez-Reif, 1998; Soken & Pick, 1992; Walker, 1982; Walker-Andrews, 1997; Walker-Andrews et al., 1991), perhaps 6-month-olds perceive similar intermodal relationships across species.

The purpose of the current experiment was to examine whether human infants perform cross-species intersensory matching of faces and vocalizations. Because the intersensory matching exhibited in Lewkowicz and Ghazanfar (2006) could be a result of infants using temporal synchrony, we chose to control for this possibility by using static rather than dynamic visual displays. We also chose to use domestic dogs (*Canis familiaris*) because dogs have been shown to display a range of facial expressions and barks that have been interpreted by some to include happiness, play, aggression, and fear (Pongrácz, Molnár, Miklósi, & Csányi, 2005). Moreover, it has also been proposed that through the process of evolution and domestication, dogs and humans have had substantial experience in communicating and interacting with one another, and therefore both species have become adept at comprehending

the various interspecific visual and auditory communicative behaviors (Cohen & Fox, 1976; Feddersen-Peterson, 2000; Miklósi, Polgárdi, Topál, & Csányi, 2000; Pongrácz, Miklósi, & Csányi, 2001; Yin, 2002). Infants between the ages of 6 and 24 months were chosen because infants of this age range are capable of matching human faces and voices using a variety of properties.

Method

Participants

One-hundred and twenty-eight infants, 32 at each of the four ages, 6, 12, 18, and 24 months, participated. The mean age of the 6-month-olds (14 girls and 18 boys) was 180.15 days. The mean age of the 12-month-olds (18 girls and 14 boys) was 366 days. The mean age of the 18-month-olds (14 girls and 18 boys) was 543 days, and for the 24-month-olds (15 girls and 17 boys), the mean age was 23 months and 29 days. Parents of the participants were initially contacted by telephone. During the initial telephone contact, parents were asked (a) whether they had a dog in the home and (a) their son or daughter's experience with or exposure to dogs. Specific to the second question, parents were asked whether their child was exposed to a dog less than 20 min a month, between 20 and 40 min month, or more than 40 min a month. Those parents who reported that (a) they did not own or have a dog in the home and (b) their child had minimal to no exposure to other dogs ($<$20 min per month) were eligible for participation in the current experiment.[1] Ninety-eight percent of the participants were White not of Hispanic origin, and 2% were Pacific Islanders. Thirty-one percent of the parents self-reported their socioeconomic status as middle–low ($<$\$35,000 per year), 58% reported their socioeconomic status as middle (\$36,000 $<$ \$50,000 per year), 9% reported their socioeconomic status as middle to upper ($>$\$50,000 per year), and 2% did not answer this demographic question. All participants were healthy, normal, full-term infants who weighed at least 5 pounds at birth.

Exclusion Criteria

The data from 15 additional infants were excluded from the study. Infants were required to complete all four trials, and the data for 7 participants (five 6-month-olds and two 12-month-olds) were rejected for failure to meet this criterion (i.e., fussiness). We also felt that it was important that infants notice that there were two video events side by side. Thus, an attention criterion required that infants look at least 10% of the time to the least preferred display. If an infant failed to meet this criterion on any of the four trials, they were excluded from the analyses. The data of 8 infants (two 12-month-olds and three 18- and three 24-month-olds) were rejected for failure to meet this criterion.

---

[1] Ideally, we wanted to compare those children who were exposed to dogs and those who had limited or no exposure to dogs. Fifteen percent of parents contacted reported owning a dog. Of those parents who reported owning a dog, they further reported that their infant was rarely exposed to the dog as the dog was often placed in kennel outside and was not permitted to enter the house. Finally, some parents ($n = 6$; across all participants) reported that they had owned a dog during their pregnancy, but upon the birth of their child, they had placed the dog for pet adoption.

Stimuli

The visual events consisted of two pairs of photographs of domestic dogs (*Canis familiaris*), and the auditory events consisted of four barks from different dogs. The barks and photographs were chosen from an initial set of 20 dogs. Dog owners were recruited from friends and colleagues of the authors. Dog owners were instructed to induce a nonaggressive or play posture and bark followed by an aggressive posture and bark. All interactions were videotaped. For each potential dog, two different static photographs were captured from the videotapes: In one image, the dog was in a neutral or nonaggressive posture, and in the other, the dog was in a negative or aggressive posture. Each pair of photos, for all 20 dogs, was rated in terms of their nonaggressive and aggressive postures and barks by 15 undergraduates using a 7-point Likert scale. The 2 dogs whose images obtained the highest overall ratings in terms of displaying an aggressive and nonaggressive or neutral posture were chosen. For the images, we chose the highest rated dog rather than the highest rated nonaggressive and aggressive picture because each participant saw the same dog side by side, and we wanted to control for any effects of having two different dogs presented side by side. Photographs used in the experiment are presented in Figure 1.

For the auditory events, in which infants only heard one bark for each in-sound trial, we chose the two highest rated aggressive and the two highest rated nonaggressive barks. We selected the first 10 s for each of the highest rated barks and looped it, creating one 20-s excerpt. The barks used in the current study are consistent with previous descriptions of aggressive and nonaggressive barks (McConnell, 1990; Yin, 2002). Specifically, aggressive barks, relative to nonaggressive barks, are more rapidly repeating so that the barks often become "fused" together or otherwise acoustically connected through growling and are of a lower pitch than nonaggressive or play barks. The barks are characterized in Figure 2. None of the barks used in the experiment were from the dogs used in the photographs.

Apparatus

The visual events were videotaped with a Sony digital camera (DSR 250; Sony Corp., Tokyo, Japan), edited with Apple's iMovie

Figure 1. Visual stimuli used in the experiment. Left-side photos show nonaggressive posture/expression. Right-side photos show aggressive posture/expression.

digital editing program (Apple Inc., Cupertino, CA), and presented with three Sony (DVP-NS57P/B) DVD players. The DVD players were connected to two 19-inch (48-cm) color video monitors (Sony KV-20M10). Soundtracks were presented from a speaker located between the monitors at approximately 65 decibels as measured from the infant's seat. The photographs were presented to the two monitors via an edit controller connected to the DVD players.

Infants sat facing the video monitors approximately 50 cm away. Two apertures cut into the black cloth surrounding the monitors were used to record the infants' visual fixations. The observers, unaware of the hypotheses of the experiment and unable to view the visual events, depressed a button while the infant fixated on the event and released it when the infant looked away. The observers also wore iPods (Apple Inc., Cupertino, CA) that played music so they would be blind to the auditory information presented to the infant. The observers' button boxes were connected to a computer programmed to record visual fixations online. A second experimenter (who controlled the presentation of the video displays) received a computer signal through a small earphone that indicated when the trials should end and when the habituation criterion had been reached. The observations of the primary observer controlled the audiovisual presentations and those of the secondary observer were used in the computation of interobserver reliability.

Counterbalancing and Procedure

The four barks were randomly divided into two pairings consisting of one aggressive and one nonaggressive bark. Half of the infants heard one pairing, and half heard the other pairing. In addition, half of the infants heard the aggressive bark first, and half heard the nonaggressive bark first. Infants saw the same dog on all four trials. Each dog was presented to half of the infants at each age. Half of the infants at each age saw the aggressive expression on the left, and half saw the nonaggressive expression on the left. The lateral position of the aggressive and nonaggressive expression did not change from trial to trial for a given infant but was counterbalanced across infants.

Infants at all ages participated in an intermodal-matching procedure (Spelke, 1976). In this and other experiments in which this procedure has been used, infants are "asked" whether they can detect the perceptual correspondence between information presented to two different sensory systems. Traditionally, infants view two visual events or images while hearing an auditory event (or in some cases haptically exploring an object) that corresponds to one of the visual events (see Kellman & Arterberry, 1998, for a review of this procedure). In many experiments using this procedure, infants look proportionally longer to the matched or congruent event (e.g., Bahrick, Hernandez-Reif, & Flom, 2005; Lewkowicz & Ghanzafar, 2006; Walker, 1982; Walker-Andrews, 1986), and in some circumstances, infants' look proportionally longer to the mismatched or incongruent event (e.g., Bahrick & Watson, 1985; Morgan & Rochat, 1995; Schmuckler & Fairhall, 2001).

During the current experiment, each infant received four 20-s trials: two silent and two in-sound trials. Only two in-sound trials were presented to reduce the possibility of infants' learning the face–bark relation during the experiment. During the two silent trials, infants viewed the static images of the aggressive and
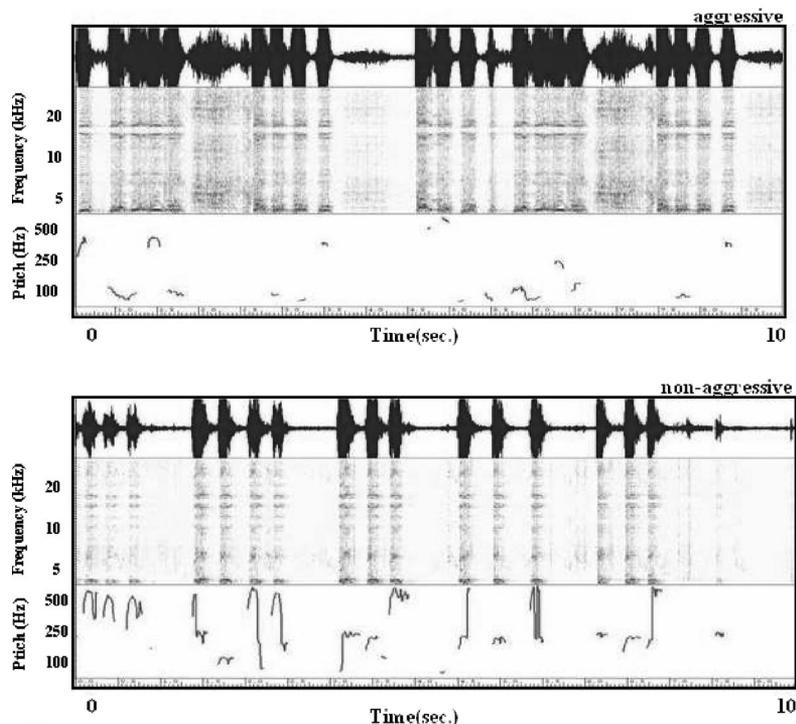
Figure 2. Acoustic properties of aggressive (upper) and nonaggressive (lower) barks. Top portion of each panel shows the acoustic oscillograms, the middle portion shows the acoustic spectrograms (measured in kilohertz), and the lower portion displays the pitch (measured in hertz).

nonaggressive expression/posture side by side. The two silent trials were always presented first. The silent trials were included to assess whether infants preferred either the aggressive or nonaggressive expression and also allowed us to examine whether infants' increased or decreased their looking to each expression when it was presented silently compared with when it was presented with its appropriate bark (see Walker, 1982; Walker-Andrews, 1986, for a similar procedure). After the two silent trials, and just prior to each in-sound trial, the bark was presented for 2 s. Following this auditory presentation, the two visual images were then presented for 20 s. During the two in-sound trials, infants heard the aggressive bark on one trial and heard the nonaggressive bark on the other trial.

Results

The primary dependent variable was the proportion of total looking time (PTLT) that infants looked toward sound/bark-specified expression. We derived proportions for each in-sound trial separately by dividing the time spent looking to the bark-specified display (i.e., congruent display) by the time spent looking at both displays (i.e., congruent + incongruent displays). We derived an overall PTLT by averaging across the two in-sound trials (one aggressive and one nonaggressive) for each infant and then averaging over all infants at each age. Infants' PTLTs to the sound-specified display are presented in Figure 3.

In order to determine whether infants demonstrated intermodal matching of the barks and expressions, we compared the mean

PTLTs against the chance value of .50 at each age. Results indicated that the 6-month-olds showed significant intermodal matching of bark and expression, $t(31) = 5.8, p < .01$, Cohen's $d = .76$ (see Figure 3). The results of the 12-, 18-, and 24-month-olds did not reach significance (all $ps > .10$).

We also performed a three-way analysis of variance with age (6, 12, 18, and 24 months), bark (aggressive, nonaggressive), and dog (Chief, Max) as between-subjects factors and infants' proportion of looking to the sound-specified image as the dependent variable. The results of this analysis revealed a significant effect of age, in which 6-month-olds' looking to the sound-specified display was greater than that of all other ages, $F(3, 112) = 4.11, p = .02, \eta_p^2 = .57$. The effect of bark, $F(1, 112) = .468, p > .1, \eta_p^2 = .03$; dog, $F(1, 112) = 2.21, p > .1, \eta_p^2 = .24$; and all interactions (all $ps > .1$) did not reach statistical significance. The fact that we did not find a significant effect of dog or Dog × Bark interaction is relevant as one dog's mouth is closed in the nonaggressive posture.

While 6-month-olds showed significant overall matching ($M = .59, SD = .09$), the data also revealed that 6-month-olds were able to match both the aggressive ($M = .61, SD = .16$) and the nonaggressive ($M = .57, SD = .13$) barks independently: $t(31) = 3.9, p < .01$, Cohen's $d = .82$, and $t(31) = 3.1, p < .01$, Cohen's $d = .80$, respectively. At no other age did infants match either the aggressive or the nonaggressive barks with the appropriate facial expression. This result shows that the significant overall intermodal matching by the 6-month-olds is not due to their matching either the aggressive or the nonaggressive
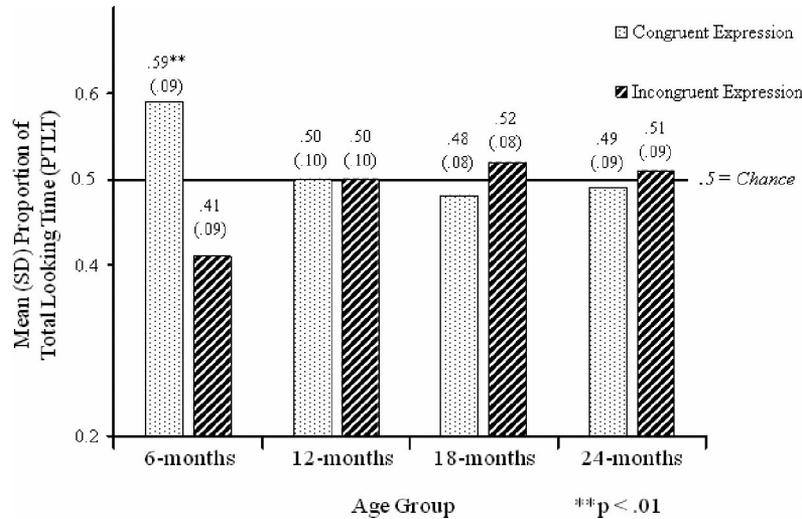
Figure 3. Mean (*SD*) proportion of total looking time (PTLT) to the bark congruent and incongruent facial expression.

bark and face; rather it shows that 6-month-olds matched both the aggressive and nonaggressive bark and face independently.

We also compared infants' proportion of time spent looking to each face in the presence of the matching bark compared with proportion of time spent looking in the absence of the bark during the silent trials. We computed two difference scores (aggressive and nonaggressive) for each infant by subtracting their proportion of time spent looking at the face in silence from the proportion of time spent looking to the face in the presence of the appropriate bark. Six-month-olds increased their looking to the aggressive expression (*M* =.12, *SD* = .200), *t*(31) = 3.3, *p* = .003, Cohen's *d* = .80, in the presence of the aggressive bark. Six-month-olds likewise increased their looking to the nonaggressive expression (*M* = .07, *SD* = .18), *t*(31) = 2.2, *p* = .037, Cohen's *d* = .74, in the presence of the nonaggressive bark. At no other age did infants significantly increase or decrease their looking to the faces in the presence of the aggressive or nonaggressive bark (all *p*s > .10). We also examined whether infants at each age differed in the PTLT for the two expressions during the two in-sound and silent trials. The results of both analyses failed to reach significance (both *p*s >.1).

In order to rule out the possibility that 6-month-olds' intermodal matching was being carried by a few infants with relatively large/strong preferences, we examined the number of infants who showed significant matching (i.e., >50%) for the aggressive and nonaggressive trials. During the aggressive in-sound trial, 24 of the 32 infants in the 6-month-old group showed significant matching, χ^2(1, *N* = 32) = 8.0, *p* < .01. Similarly, during the nonaggressive in-sound trial, 23 of the 32 infants in the 6-month-old group showed significant matching, χ^2(1, *N* = 32) = 6.3, *p* = .013. Thus, the results of the 6-month-olds are likely not due to a few infants with large preferences.

Taken together, the results of the 6-month-olds show reliable matching for both the aggressive and nonaggressive barks, whereas the results of the older infants are more ambiguous because they did not show reliable matching for the aggressive

or nonaggressive bark. One possibility is that older infants indeed matched the appropriate expression and bark during the initial portion of the trial and then switched their gaze to the incongruent display. In order to assess this possibility, we examined infants' looking to the matching expression separately during the first 10 s and the second 10 s for the two in-sound test trials. Infants' PTLT to the sound-specified display (i.e., congruent expression) for the first and second 10 s are presented in Figure 4.

Results indicate that the 6-month-olds looked proportionately longer to the matching or congruent display during the first as well as the second 10 s of the test trials, *t*(31) = 5.6, *p* < .01, Cohen's *d* = .92, and *t*(31) = 5.3, *p* < .01, Cohen's *d* = .90, respectively. Results of the 12-month-olds, however, failed to show reliable matching, or mismatching, during the first or second 10 s of the test trials (both *p*s > .1). In contrast to the 6- and the 12-month-olds, results of the 18- and 24-month-olds during the second 10 s of the test trials show a significant preference for the incongruent display, *t*(31) = 4.4 *p* < .01 Cohen's *d* = .84, and *t*(31) = 2.3, *p* < .05, Cohen's *d* = .67, for the 18- and 24-month-olds, respectively. In addition to comparing infants' looking during the first and second 10 s of the in-sound trials, we also examined the direction of infants' first look during these two trials. Infants' first look was coded as congruent if it was directed toward the appropriate monitor and was at least .5 s in duration. Table 1 shows the number of infants at each age for each in-sound trial whose first look was directed toward the sound-specified face.

Results for the 6-month-olds reveal that the number of infants whose first, or initial, look was directed toward the correct or matching face did not reach significance on nonparametric chi square tests (both *p*s > .1). Results for the 12-month-olds reveal that 23 of the 32 infants' (71%) first look during the second test trial was toward the matching face, and this number reached statistical significance, χ^2(1, *N* = 32) = 6.13, *p* < .05. Results for the 18- and 24-month-olds also indicate that the number of infants whose first look during the first trial was directed
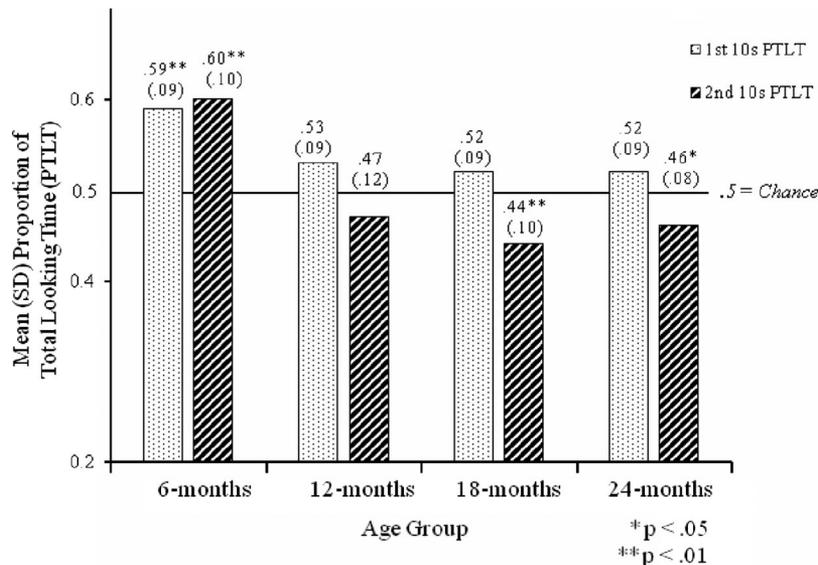
Figure 4. Mean (*SD*) proportion of total looking time (PTLT) to the bark congruent expression during the first and second 10s of the test trials.

toward the matching display exceeds chance, $\chi^2(1, N = 32) = 12.5$, $p < .01$, and $\chi^2(1, N = 32) = 15.1$, $p < .01$, respectively. The number of 18- and 24-month-olds whose first look during the second trial was directed toward the congruent face also exceeds chance, $\chi^2(1, N = 32) = 8.0$, $p < .01$, and $\chi^2(1, N = 32) = 6.13$, $p < .05$.

Taken together, these results show overall matching of aggressive and nonaggressive canine barks and facial expressions by infants at 6 months of age. Six-month-olds also increased their looking to aggressive and nonaggressive expressions in the presence of the congruent bark. Results also demonstrate that older infants' initial looks were directed toward the appropriate expression, and during the latter half of the test trials, older infants looked proportionately longer to the incongruent expression.

Discussion

The results of this study are the first to show that infants perform cross-species intersensory matching of canine facial expressions and barks. In general, these results parallel those of Lewkowicz and Ghazanfar (2006) as younger, but not older, infants looked proportionately longer to the congruent expres-

Table 1
Number of Infants in Each Age Group Whose First Look Was Directed Toward the Sound-Specified Face

	First look			
Trial	6 months	12 months	18 months	24 months
First	18	21	26**	27**
Second	20	23*	24**	23*

Note. Number of infants in each of the four age groups was 32.
* $p < .05$.   ** $p < .01$.

sion. These results also extend those of Lewkowicz and Ghazanfar (2006) as 6-month-olds demonstrated cross-species intersensory matching in the absence of temporal synchrony. In addition, these results also reveal that in regard to infants' first looks and comparisons of infants' proportion of looking during the first and second half of test trials, older infants showed evidence of cross-species intermodal matching. These results also suggest that cross-species intersensory perception (at least for canine faces and vocalizations) does not decline and that singular measures such as infants' PTLT may be insufficient at capturing this ability across a wide range of ages.

In terms of determining on what basis infants are making the cross-species intersensory match, one possibility is that 6-month-olds perceive the common affective information from the barks and the facial expressions. The results from the current experiment, as well as from other experiments, support such a possibility. With human faces and voices, for example, infants between 5 and 7 months of age match angry and happy faces and voices (Walker, 1982). Seven-month-olds also make similar intersensory matches even when the information for face–voice synchrony is absent and when facial information is conveyed using only point lights (Soken, & Pick, 1992; Walker-Andrews, 1986). Moreover, the acoustic properties of the aggressive and nonaggressive barks are consistent with those properties of aggressive and nonaggressive speech in humans (Fernald, 1989; Pongrácz et al., 2005; Williams & Stevens, 1972). In addition to the acoustic similarities, there is evidence that the aggressive facial expressions of canines share some features with angry human faces (Aronoff, Barclay, & Stevenson, 1988; Darwin, 1898/2007; Ekman & Friesen, 1975; Gross, 2004; Lundqvist, Esteves, & Öhman, 1999; McKelvie, 1973).

In terms of determining whether cross-species intersensory perception declines with age, Lewkowicz & Ghazanfar (2006) found that younger (4- and 6-month-olds), but not older (8- and 10-

month-olds), infants looked proportionately longer at the congruent expression; however, they did not examine the direction of infants' first look nor did they examine whether infants' looking behavior differed as the test trial progressed. The fact that we found older, but not younger, infants' frequency of first looks to be in the direction of the congruent expression and that older infants also looked significantly longer to the incongruent expression during the latter half of the test trial suggests cross-species intersensory perception may not decline. Moreover, the results of the older infants suggest that they were able to quickly discern which facial or postural expression corresponded to the bark and then focused their visual attention toward the incongruent expression. In contrast, the number of first looks for the 6-month-olds that were directed toward the congruent expression, while slightly greater than 50%, failed to reach significance. Still 6-month-olds did look proportionately longer, during both the first half and second of the test trials, to the congruent expression. Finally, there is also evidence that adults, including dog owners as well as nonowners, discriminate and categorize dog barks as aggressive or playful (Pongrácz, et al., 2005), which further suggests this ability does not decline. Taken together, it seems infants between 6 and 24 months of age perceive cross-species intersensory relationships for canine faces and barks; however, the nature with which infants' exhibit this intersensory matching varies across age and therefore requires different measures (i.e., first looks, proportion of total looking time).

Just as there are differences between the current study and the work of Lewkowicz and Ghanzafar (2006) in terms of what looking behavior was coded, there are also differences in the auditory and visual events. The most obvious difference is that the visual events in Lewkowicz and Ghanzafar's (2006) study were dynamic, whereas we chose static visual displays. While dynamic displays can, in a more ecologically valid way, convey communicative information relative to static displays, two communicatively different dynamic visual displays cannot be temporally aligned to a single auditory event. Again, because we were interested in examining whether infants could perceive the cross-species intersensory relationship without using temporal synchrony, we chose static images. In addition, rhesus macaques, like canines and other organisms, communicate through posture or body position, presence or absence of teeth, positioning of ears, and having their body hair stand on end (e.g., Hauser, 1996; Jolly, 1985; Nystrom & Ashmore, 2008). Thus, we chose to use photographs that showed each canine's head as well as some portion of the dog's body position or posture, whereas Lewkowicz and Ghazanfar (2006) presented only the head of the monkey. Finally, it has also been proposed that through the process of evolution and domestication, humans and dogs have become adept at comprehending the various interspecific visual and auditory communicative behaviors (Cohen & Fox, 1976; Feddersen-Peterson, 2000; Miklósi et al., 2000; Pongrácz et al., 2001; Yin, 2002). It is possible, therefore, that the different patterns of results obtained in the current experiment, compared with the results of Lewkowicz and Ghazanfar (2006), may in part be due differences in the species used in each experiment as well as the nature of the visual events.

In a broader context, these results, as well as those of Lewkowicz and Ghazanfar (2006), begin to address the question of how intersensory perception changes over time. Many studies of intersensory perception have used human faces and voices, for example, to examine when this perceptual ability emerges, but little to no evidence is available concerning whether and how this ability changes over time. In contrast, research examining unimodal visual or unimodal auditory perception has documented how infants' discrimination of faces and speech phonemes emerges and also changes over time and experience (e.g., Kuhl et al., 1992; Pascalis et al., 2002; Pascalis et al., 2005 Werker & Tees, 1984). In addition, research in which the development of memory has been examined with the preferential looking procedure has shown infants' visual preference is dynamic. Specifically, infants' visual recognition memory is initially expressed as a novelty preference, changes to a null preference, and then changes to a familiarity preference as memory accessibility declines across retention time (Bahrick, Gogate, & Ruiz, 2002; Bahrick, Hernandez-Reif, & Pickens, 1997; Bahrick & Pickens, 1995). Therefore, only in the areas of unimodal auditory and visual perception, as well as infants' memory for visual events, have researchers explored how these abilities develop and change over time; research examining how intersensory perception changes over time is needed.

One possibility for future research includes examination of whether and how older infants' intersensory perception for affect, as conveyed in human faces–voices, changes over time. It is well established that this ability emerges by the age of 5–7-months, yet it is unknown whether and how this ability changes (Walker-Andrews, 1997). Perhaps, as the current results suggest, it is possible that younger infants look longer to congruent affective facial expression, whereas older infants look longer to the incongruent affective facial expression. Additional research conducted with convergent measures is needed to explain how intersensory perception develops and changes over time.

If one function or purpose of intersensory perception is to guide early learning, then, as suggested by Lewkowicz and Ghazanfar (2006), it may be beneficial to the young infant to have a broad intersensory–perceptual system in order to capture as many intersensory relationships as possible. Because much early learning occurs within the context of interacting and communicating with others, where the sharing of affect is one of the first forms of communication, it may be advantageous for infants to perceive intersensory matches for affect at an early age and across a variety of faces and voices. In addition, just as research has examined how infants' unimodal perception of faces and voices changes over the course of development, research should examine how infants' intermodal perception changes during development. Finally, cross-species studies of face perception and, now, intersensory perception highlight (a) the fact that infants are sensitive to a variety of communicative behaviors and (b) the experience-dependent nature of early perceptual development.

References

Aronoff, J., Barclay, A. M., & Stevenson, L. A. (1988). The recognition of threatening facial stimuli. *Journal of Personality and Social Psychology, 54,* 647–655.

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.

Bahrick, L. E., Hernandez-Reif, M., & Pickens, J. N. (1997). The effect of retrieval cues on visual preferences and memory: Evidence for a four-phase attention function. *Journal of Experimental Child Psychology, 67,* 1–20.

Bahrick, L. E., Hernandez-Reif, M. R., & Flom, R. (2005). The development of infant learning about specific face-voice relations. *Developmental Psychology, 41,* 541–552.

Bahrick, L. E., Netto, D., & Hernandez-Reif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Development, 69,* 1263–1275.

Bahrick, L. E., & Pickens, J. N. (1995). Infant memory for object motion across a period of three-months: Implications for a four-phase attention function. *Journal of Experimental Child Psychology, 59,* 343–371.

Bahrick, L. E., & Watson, J. S. (1985). Detection of intermodal proprioceptive–visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology, 21,* 963–973.

Barrera, M., & Mauer, D. (1981). Recognition of mother's photographed face by the three month old infant. *Child Development, 52,* 714–716.

Birnholtz, J. C., & Benacerraf, B. R. (1983, November 4). The development of human fetal hearing. *Science, 222,* 516–518.

Cohen, J. A., & Fox, M. W. (1976). Vocalizations in wild canids and possible effects of domestication. *Behavioural Processes, 1,* 77–92.

Darwin, C. (2007). *The expression of the emotions in man and animals.* (3rd ed.). Stuttgart, Germany: Schweizerbart. (Original work published 1898)

DeCasper, A. J., & Fifer, W. (1980, June 6). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174–1176.

DeCasper, A. J., & Spence, M. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9,* 133–150.

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology, 11,* 478–484.

Easterbrook, M., Kisilevsky, B., Hains, S., & Muir, D. (1999). Faceness or complexity: Evidence from newborn visual tracking of facelike stimuli. *Infant Behavior and Development, 22,* 17–35.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotion from facial clues.* Englewood Cliffs, NJ: Prentice-Hall.

Fagan, J. F. (1972). Infants' recognition memory for faces. *Journal of Experimental Child Psychology, 14,* 453–476.

Feddersen-Peterson, D. U. (2000). Vocalization of European wolves (*Canis lupus lupus L.*) and various dog breeds (*Canis lupus f. fam*). *Archiv für Tierz Dummerstorf, 43,* 387–397.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60,* 1497–1510.

Gross, T. F. (2004). The perception of four basic emotions in human and nonhuman faces by children with autism and other developmental disabilities. *Journal of Abnormal Child Psychology, 32,* 469–480.

Hauser, M. D. (1996). *The evolution of communication.* Cambridge, MA: Bradford/MIT Press.

Jolly, A. (1985). *The evolution of primate behavior* (2nd ed.). New York: Macmillan.

Kellman, P. J., & Arterberry, M. (1998). *The cradle of knowledge: Perceptual development in infancy.* Cambridge, MA: MIT Press.

Kleiner, K. A. (1987). Amplitude and phase spectra as indices of infants' pattern preferences in human infants. *Journal of Experimental Child Psychology, 7,* 339–354.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development, 7,* 361–381.

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. *Proceedings of the National Academy of Sciences of the United States of America, 100,* 9096–9101.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992, January 31). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 6771–6774.

Lundqvist, D. L., Esteves, F., & Öhman, A. (1999). The face of wrath: Critical features for conveying facial threat. *Cognition & Emotion, 13,* 691–711.

Mauer, D., & Young, R. E. (1983). Newborns' following of natural and distorted arrangements of facial features. *Infant Behavior and Development, 6,* 127–131.

McConnell, P. B. (1990). Acoustic structure and receiver response in domestic dogs *Canis familiaris. Animal Behaviour, 39,* 897–904.

McKelvie, S. J. (1973). The meaningfulness and meaning of schematic faces. *Perception and Psychophysics, 14,* 343–348.

Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (2000). Intentional behavior in dog–human communication: An experimental analysis of "showing" behavior in the dog. *Animal Cognition, 3,* 159–166.

Morgan, R., & Rochat, P. (1995). Spatial determinants in the perception of self-produced leg movements by 3- to 5-month-old infants. *Developmental Psychology, 31,* 626–636.

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development, 10,* 3–18.

Nystrom, P., & Ashmore, P. (2008). *The life of primates.* Upper Saddle River, NJ: Pearson Prentice–Hall.

Pascalis, O., de Hann, M., & Nelson, C. A. (2002, May 17). Is face processing species-specific during the first year of life? *Science, 296,* 1321–1323.

Pascalis, O., & de Schonen, S. (1994). Recognition in 3- to 4-day-old human neonates. *NeuroReport, 5,* 1721–1724.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 5297–5300.

Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development, 22,* 237–247.

Pongrácz, P., Miklósi, Á., & Csányi, V. (2001). Owners' beliefs on the ability of their pet dogs to understand human verbal communication. A case of social understanding. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 20,* 87–107.

Pongrácz, P., Molnár, C., Miklósi, Á., & Csányi, V. (2005). Human listeners are able to classify dog (*Canis familiaris)* barks recorded in different situations. *Journal of Comparative Psychology, 119,* 136–144.

Querleu, D., Renard, X., Boutteville, C., & Crepin, G. (1989). Hearing by the human fetus? *Seminars in Perinatology, 13,* 409–420.

Quinn, P. C., Yarr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31,* 1109–1121.

Schmuckler, M. A., & Fairhall, J. L. (2001). Infants' visual–proprioceptive intermodal perception of point-light information. *Child Development, 72,* 949–962.

Soken, N. H., & Pick, A. D. (1992). Infants' perception of dynamic affective expressions: Do infants distinguish specific expressions? *Child Development, 70,* 1275–1282.

Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8,* 553–560.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology, 13,* 514–535.

Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology, 22,* 373–377.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin, 121,* 437–456.

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender, *Ecological Psychology, 3,* 55–75.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *The Journal of the Acoustical Society of America, 52,* 1238–1250.

Yin, S. (2002). A new perspective on barking in dogs (*Canis familiaris*). *Journal of Comparative Psychology, 119,* 189–193.