I don't like the tone of your voice: Infants use vocal affect to socially evaluate others

Melissa Paquette-Smith and Elizabeth K. Johnson

Department of Psychology, University of Toronto

3359 Mississauga Road North, Mississauga, ON, L5L 1C6, Canada

Author Note

We thank actors Samantha Piper, Miranda Jary, RJ Mojica and David Ball for recording the speech stimuli for this study. We thank Alyssandra Chee-A-Tow for help with Figure 1. NSERC and SSHRC grants awarded to EKJ supported this work. Part of this work was presented at the 18th International Conference on Infant Studies, Minneapolis, MN and at the 2011 Society for Research in Child Development Biennial Meeting, Montreal, QC.

Correspondence should be addressed to Melissa Paquette-Smith, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road North, Mississauga, Ontario, Canada, L5L 1C6. Email: m.paquette.smith@mail.utoronto.ca. Phone: 905-569-4859. Fax: 905-569-4850.

Keywords: tone of voice; speech perception; emotional affect

Word count: 6,336

Revision Submitted: June 16, 2015

Abstract

Infants can make social judgments about characters by visually observing their interactions with others (e.g., Hamlin, Wynn & Bloom, 2007). Here, we ask whether infants can form similar judgments about potential social partners based solely on their tone of voice. In Experiment 1, we presented 10.5-month-olds with two visually neutral puppets. One puppet spoke in a positive affect and the other spoke in a negative affect. When the puppets were placed within reach of the infants, infants selected the formerly positive puppet. This preference disappeared when the voices were paired with non-social objects (Experiment 2). In Experiment 3, 10.5-month-olds were once again exposed to the same emotionally negative and positive voices. However, no visual characters were present. At test, infants' visual orientation controlled how long they heard the neutral versions of each voice. Here, infants listened longer to the neutral voice of the formerly positive speaker. That is, just as in Experiment 1, infants' preferences for the emotionally neutral test stimuli was shaped by their earlier exposure to emotionally charged recordings of that voice. Our findings provide convergent evidence to suggest that infants possess sophisticated social evaluation abilities, preferring to interact with prosocial over antisocial others.

Keywords: tone of voice; speech perception; emotional affect

Word Count: 196

I don't like the tone of your voice: Infants use vocal affect to socially evaluate others

The ability to evaluate the actions and intentions of others is an important skill that enables us to respond adaptively to social situations. The average adult possesses highly advanced social evaluation skills, making snap judgments on the character of others in as little as a few seconds (Ambady & Rosenthal, 1993). Children too can use visual information to make social judgments about others relatively quickly (e.g., Hamlin, Wynn & Bloom, 2007). These skills are important to navigate the social world, as is evidenced by the difficulties experienced by individuals with autism who struggle to interpret social information on the faces and voices of others (Hobson, Ouston & Lee, 1998). However, many aspects of how infants evaluate potential social partners remain unexplored. For example, it is still unclear how sensitive young children are to cues to vocal emotion. In this paper, we begin to address this issue by examining 10.5-month-olds' reaction to vocal affect information presented in different social scenarios.

In the past 20 years we have learned that infants possess surprisingly sophisticated social skills. By the end of the first year of life, typically developing infants can classify visual displays of social acts as being positive or negative (Premack & Premack, 1997). Moreover, infants can use these overt displays of behavior to help them select appropriate social partners. After watching a scenario in which a character is helped by one individual and hindered by another, 6-month-olds will choose to interact with the helping character over the hindering character (Hamlin et al., 2007). A few months later, infants are able to use this information to predict how the character would likely respond to individuals that helped or hindered them (Fawcett & Liszkowski, 2012).

Although studies have demonstrated that infants can evaluate others based on their visually observable actions, not all social behaviors can be observed through the visual modality.

For example, some social evaluations can be made simply by over-hearing a person speak. In scenarios such as this where visually observable behaviors are unavailable, it would be advantageous to be able to evaluate others based on information coming through the auditory channel. At present, we only have a limited understanding of children's ability to interpret vocal affect cues. There is evidence that vocal affects (i.e., approvals and disapprovals) are universally understood across languages and cultures, which could suggest that the ability to interpret this information is innate (Bryant & Barrett, 2007). Early on in development, infants have been shown to be highly responsive to auditory social cues. For example, 3-month-olds can differentiate between positively and negatively affected voices (Walker-Andrews & Grolnick, 1983). And by 7 months of age, infants show an understanding of these vocal affects by orienting quickly to negative or threatening voices (Grossmann, Striano & Friederici, 2005), while at the same time preferring to listen to positive over negative or neutral voices (Singh, Morgan & Best, 2002; Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990). There is even evidence that older infants can use auditory social cues (such as fearful vocalizations) to regulate their behavior (i.e., to decide whether to approach or avoid a novel toy; Mumme, Fernald, & Herrera, 1996; Vaish & Striano, 2004). Although these studies demonstrate that infants are responsive to the emotional valance of different vocal affects, we do not yet know the extent to which they can retain and utilize the speaker's vocal affect to guide their future social encounters with that individual.

There is evidence that infants can use certain types of auditory cues to evaluate potential social partners. For example, infants prefer characters who speak in a familiar language over those who speak in an unfamiliar language (Kinzler, Dupoux, & Spelke, 2007). But no work to date has examined whether infants can use a speaker's tone of voice in a similar way. Existing

research suggests that even school-age children can have difficulty interpreting tone of voice cues in certain contexts (Morton & Trehub, 2001; Morton, Trehub, & Zelazo, 2003). However, that many of these studies pit the semantic content of the utterance against the vocal affect expressed by the speaker. For example, a speaker might say 'my dog ran away from home' in a positive affect or 'I got an ice cream for being good' in a negative affect (Morton & Trehub, 2001). The ecological validity of such a task can be questioned since children are not routinely faced with utterances of this sort. Moreover, the task of interpreting spoken utterances that contain mismatching semantic content and vocal affect likely assesses a different level of emotional understanding than that needed for social evaluation. Given the precocious nature of infants' use of visually observable behaviors to make social judgments, it is possible that if presented with a simplified auditory task, they may show a better understanding of vocal affect than previous studies have demonstrated (see Berman, Chambers, & Graham, 2010 for related discussions).

Here, we present three experiments examining infants' use of vocal affect cues to socially evaluate others. In Experiment 1, after observing puppets speaking in a positive or negative tone of voice, 10.5-month-olds' are allowed to select a puppet to interact with. Next, in Experiment 2, we report a control study designed to rule out the possibility that infants' behavior in Experiment 1 was driven by simple non-social associations. And finally, in Experiment 3 we test whether 10.5-month-olds' can make similar social judgments without the aid of any visual information. In this last experiment, infants were exposed to a positive and a negative voice, and then allowed to control the presentation of those speakers' neutral voices during the test phase. Importantly, all test stimuli were neutral in all experiments, thus any preferences demonstrated by the infants must be driven by the judgments they formed in the exposure phase, when the characters spoke

in either a positive or a negative affect. In combination, the results of all three experiments provide convergent evidence that infants are capable of using vocal affect information to make social judgments.

Experiment 1

In this experiment, infants watched as a positive and a negative affected puppet took turns speaking. Immediately after hearing both puppets in the pair speak infants were prompted to select a puppet to 'play with'. Each child watched four different puppet exposures, two with the first pair of puppets and voices and two with the second pair. We predicted that if infants could use the puppets' vocal affect to evaluate them as being appealing or aversive, then they should select the puppets they had seen speaking in a positive tone of voice.

Method

Participants. Twenty Canadian English learning infants between 10.5- and 11.5-months of age (M_{Age} = 341.6 days; range 323-350 days; 11 females) were tested. Data from five additional infants were excluded from the study. Four due to fussiness, and one because they failed to ever make a selection.

Stimuli. Two pairs of male and female native English speakers were recorded producing 8 semantically neutral sentences in positive (happy) and negative (irritated) vocal affects (see Appendix A). To ensure that the positive and negative sentences conveyed the intended emotion, we had 10 adults (M_{age} = 25; 6 females) rate each passage on a scale from 1 (positive/happy) to 9 (negative/irritated), with 5 being neutral. Although all of the sentences were semantically neutral (see Appendix A) the stimuli was low pass filtered to ensure that semantics did not influence the adults' ratings. Adults rated the passages as falling into two distinct emotional categories, with the positive passages being rated as closer to 1 (positive/happy) (M = 2.92, SD = 0.57) than the

negative passages (M = 6.74, SD = 0.56); $t(9) = 13.95$, $p <$. 001. Similar to Singh et al. (2002) adults were also asked to rate the low-pass filtered passages on their emotional intensity on a scale from (1 emotionally flat/neutral) to (7 intensely emotional). Here we found no statistically significant differences in intensity between the positive (M = 5.24, SD = 0.15) and the negative passages (M = 5.01, SD = 0.52), $t(10) = 1.54$, p = .16, making it unlikely that differences in intensity could drive infants' selections.

The puppet show was performed live by pairing the positive passages from one speaker with one puppet and the negative passages from the other speaker with the other puppet. There were two sets of two puppets. All four puppets had the same neutral facial expression and were closely matched in size and complexity (see Figure 1).

Design. The experiment consisted of two blocks of two trials. The first block of trials used one set of puppets and male/female voices, and the second block used the second set. In each trial the puppets alternated so that one puppet would say a pair of sentences in one affect (e.g., positive), then the other puppet would say the same pair of sentences in the other affect (e.g., negative). In half the trials the negative puppet spoke first and in the other half the positive puppet spoke first. The trial ended after each puppet had spoken twice, and the child was given the opportunity to select. Although each infant only viewed one puppet in each block as positive and the other as negative, the emotional affects (negative or positive) of the puppet as well as the side of the display the puppet appeared on were counterbalanced across participants.

Procedure. Infants sat on their caregivers lap across from the display (see Figure 1). Parents were asked to listen to masking music over headphones. The masking music consisted of loud continuous music mixed with multiple overlaid tracks of the speech stimuli used in the study. The parents could not hear the emotional affects their child was listening to, so they could

not cue them to select a particular character. The experimenter was hidden behind a curtain at all times, and could not see or be seen by the child.

In the exposure phase, the experimenter opened and closed the puppet's mouths in synchrony with the auditory passages. To give illusion that the auditory affects were coming from the puppets, the passages played from audio speakers located behind each puppet on the right or left sides of the display. After each 1-minute puppet exposure the experimenter removed their hands from inside the puppets and the puppets were re-attached to the puppet board using magnets. The parent was instructed to position the child's hands on top of two handprint markings equidistant from the left and right puppets. This was to ensure the infant would have equal opportunity to grab either puppet when they were in reach. Once the parent positioned and then released the child's hands, the experimenter slid the display across the table towards the child. The puppet display was attached to a set of wooden guides that ensured that both puppets would arrive at the same rate and would be equidistant from the child's hands (see Figure 1). Once the puppets were within reach of the child the experimenter would call out "'*Child's Name*' which one would you like to play with?" Generally infants' made their selections quickly (Mean time to select = 10.31s). If the child failed to select after 90s (1.5 minutes) the experimenter moved on to the next trial without recording a selection. Selections were videotaped and coded offline. All selections were re-coded by a second coder and inter-rater reliability was high (Kappa = .98, $p < .0001$). Trials in which infants grabbed both puppets off the board simultaneously or failed to select were excluded.

Results and Discussion

Thirteen out of 20 infants selected a puppet on all four trials (see Experiment 1, Figure 2). The remaining children picked one puppet on at least two trials (Range 2-4 selections; Mean total

number of selections = 3.6). We based our analysis on the proportion of correct selections, which was calculated by dividing the number of times the formerly positive puppet was selected by the total number of selections. For infants that made selections in both blocks, there were no differences between their performance in the first block (which used the first pair of voices and puppets; M = .63, SD = .37) and the second block (which used the second set of voices and puppets; M = .68, SD = .38), $t(18) = -0.52$, p = .61, thus for this analysis the portion correct was collapsed across both blocks. When the portion of correct selections was compared to chance (.5), the results of a two-tailed one sample t-test indicated that infants were more likely to select the puppet that had spoken with a positive tone of voice (M = .67, SD = .30), $t(19) = 2.46$, p = .02, d = .55. The speakers, their affects, and side of the display the puppet appeared on were counterbalanced across participants. Overall, infants showed no preferences for puppets that were paired with female speakers over male speakers (M_{female} = .48, SD = .24), $t(19) = -.31$, p = .76, or for puppets that appeared on the right side of the display over the left side ($M_{right\ side}$ = .54, SD = .35), $t(19) = .54$, p = .60.

The results of Experiment 1 suggest that 10.5-month olds infants can evaluate novel social partners based on their tone of voice. Previous research has shown that infants prefer to listen to positive over negative voices (Papoušek et al., 1990), however, this is the first study to suggest that infants can use this information to make a judgment about the social desirability of the speaker. However, although infants' selections in this sort of experiment are typically interpreted as the product of a complex social evaluation process (Hamlin et al., 2007; Kinzler et al., 2007), we readily acknowledge that there may be other lower-level explanations for our results. Most obviously, instead of socially evaluating the puppets infants might have simply selected the puppet that they perceptually associated with the positive voice. It is even possible

that infants were simply associating the positive or negative tone of voice with the left or right side of the stage rather than attaching it to the puppets (each puppet consistently appeared on the same side of the display throughout the puppet show and selection phases - the characters did not switch sides during the experiment). Although both of these alternative explanations require that the child retain and use vocal affect information, they do not support the notion that children's choices reflect a social judgment. Experiment 2 is designed to examine these alternative explanations by replacing the puppets with non-social objects.

Experiment 2

Although it is possible that infants' selections in Experiment 1 are a reflection of their sophisticated social evaluation abilities, it is also possible that infants' behavior was simply driven by associating the puppets or the side of the display with the positive/negative vocalizations. Experiment 2 was designed to rule out these associative explanations. Similar to Hamlin et al. (2007), in our non-social control we replaced the animate puppets with inanimate objects (children's sippy cups). The non-social sippy cups moved in synchrony with the auditory passages but had no eyes, or facial features that could indicate animacy. Here we predicted that if the results of Experiment 1 are evidence of social evaluation rather than simple associations, then infants should show no preference for the sippy cup paired with the positive voice over the sippy cup paired with the negative voice.

Method

Participants. Twenty Canadian English learning infants between 10.5- and 11.5-months-old (M_{Age} = 330.85 days; range 316 – 351 days; 8 females) were tested. The data from three additional children were excluded from the analysis due to fussiness (1), experimenter error (1)

and because the child failed to ever make a selection (1).

Stimuli. All auditory stimuli were identical to Experiment 1. Instead of puppets, the positive and negative passages were paired with empty spill-proof sippy cups. The sippy cups were matched in shape and size, but varied in color.

Design and Procedure. The experimental design and procedure were identical to Experiment 1 except for the replacement of the puppets with sippy cups. During the exposure phase, the passages played from behind the sippy cups and the experimenter used a stick attached to the back of each cup to moved it up and down in synchrony with the audio passage. At the end of the exposure phase the sippy cups were placed into slots on the base of the display and the display was pushed towards the child for selection. Overall, infants' selected quickly (Mean time to select = 6.43s). If the child failed to select after 90s (1.5 minutes) the experimenter was instructed to move on to the next trial without recording a selection. All selections in this experiment were re-coded by a second coder and inter-rater reliability was high (Kappa = .93, $p < .0001$). Trials in which infants grabbed both sippy cups simultaneously or failed to select after 90s (1.5 minutes) were excluded.

Results and Discussion

Fifteen out of 20 infants selected a sippy cup on all four trials (see Experiment 2, Figure 2). The remaining children picked on at least two trials (Range 3-4 selections; Mean total number of selections = 3.8). There were no differences between infants' performance in the first block (which used the first pair of voices and sippy cups; $M = .48$, $SD = .44$) and the second block (which used the second set of voices and sippy cups; $M = .50$, $SD = .40$), $t(19) = -0.25$, $p = .80$. Thus, as in Experiment 1, the proportion of correct selections was calculated by dividing the number of times the formerly positive sippy cup was selected by the total number of selections.

Infants selections were at chance (M = .48, SD = .36), $t(19) = -0.21$, p = .84, indicating that they had no preference for the sippy cups paired with the positive or the negative speakers. As in Experiment 1, infants showed no preferences for female over male puppets (M_{female} = .55, SD = .23), $t(19) = .99$, p = .33, and for puppets positioned on the right versus the left side of the display ($M_{right\ side}$ = .52, SD = .36), $t(19) = .21$, p = .84.

Experiment 2 controlled for several alternative explanations for the results of Experiment 1, including the possibility that selections might be driven by a non-social association between the puppets and the positive/negative vocalizations. Here, when the emotional speakers from Experiment 1 were paired with inanimate sippy cups, infants showed no preference for the sippy cup paired with the nice speaker. Although this is an indication that the selections seen in the first experiment are likely driven by social rather than perceptual mechanisms, it is also possible that differences in attention to the social vs. the non-social stimuli could have impacted infants' ability to remember the voice-object pairings. As in other studies demonstrating greater attention to face-like stimuli (e.g., Fantz, 1964; Frank, Vul, & Johnson, 2009; see Pascalis & Kelly, 2009 for a review), infants spent a greater proportion of the trial looking at the social puppets in Experiment 1 (M = .82, SD = .12) compared to the non-social sippy cups in Experiment 2 (M = .67, SD = .11), $t(36) = 4.02$, $p < .001$. Although we cannot rule out the possibility that decreases in attention may have affected children's ability to remember to sippy cup-voice pairings in the control experiment, it seems unlikely given that in both experiments infants spending majority of the trial orientating towards the visual display. To address these alternative explanations, our third experiment will examine infants' ability to evaluate speakers in the absence of any visual aids.

Experiment 3

In Experiment 3, we ask whether infants can still evaluate the nature of the speakers if they are only exposed to their disembodied voices speaking in a positive or negative affect. Unlike the first two experiments, Experiment 3 is an entirely auditory paradigm. Here we exposed 10.5-month-old infants to one happy voice and one irritated voice. Then we tested infants' preference for the speakers' neutral voices using an infant-controlled testing paradigm in which the infants' looking behavior determined how long a speakers' voice was presented.

Method

Participants. Sixteen Canadian English learning infants between 10.5- and 11.5-months-old (M_{Age} = 336.25 days; range 324 – 351 days; 8 females) were tested. The data from one child was excluded due to fussiness.

Stimuli. Experiment 3 used the positive and negative audio recordings from one of the two sets of male-female speaker pairs presented in Experiments 1 and 2. The same male and female speaker also recorded neutral versions of the 8 sentences. To ensure that the neutral sentences were emotionally equidistant from the positive and negative sentences, we had 11 adults (M_{age} = 19; 9 females) who were blind to the hypothesis of the study rate the sentences on a scale from 1 (positive/happy) to 9 (negative/irritated), with 5 being neutral. The result of a repeated measures ANOVA with speaker (1 or 2) and affect (happy, irritated, neutral) indicated a significant effect of affect, $F(1.29, 12.93) = 249.57$, $p < .001$, $\eta_p^2 = .96$, and no effect of speaker $F(1, 10) = 0.21$, $p = .66$, $\eta_p^2 = .02$. As was expected, the adults' ratings fell into three distinct emotional categories, positive ($M = 2.33$, $SD = 0.67$), neutral ($M = 4.83$, $SD = 0.28$) and negative ($M = 7.71$, $SD = 0.51$; see Figure 3). Thus, we concluded that the stimuli were well suited for use in our infant study.

Design. In the exposure phase, each infant listened to a 2-minute passage of a positive affected speaker and a negative affected speaker. Although each infant only heard one speaker as positive and the other as negative, the order of the passages and the vocal affects of the speakers were counterbalanced across participants. To ensure equal exposure to each voice, the recordings of the two speakers alternated during the exposure phase. The test phase consisted of two blocks of four trials. In each block infants were presented with two trials presenting the neutral voice of the male speaker and two trials presenting the neutral recordings of the female speaker.

Procedure. This experiment used a variant of the Headturn Preference Procedure in which the infant was seated on their caregiver's lap in the middle of a set of three lights (see Johnson & Zamuner, 2010). Parents wore headphones with masking music to avoid them from biasing their child's responses. Each infant was exposed to a 2-minute passage of a positive affected speaker and a negative affected speaker. In the exposure phase, the lights blinked on and off depending on where the child looked, but the exposure passage played continuously from audio speakers located on the left and right sides of the booth. In the test phase the duration of the lights and the audio passage was contingent on the infant's looking behavior. At the beginning of each trial the center light flashed to get the infant's attention. Once they oriented towards the center light, the left or right light would begin to flash. The side that the light appeared on (left or right) was computer randomized, thus there should be no contingency between the voice and the side lights. Once the infant oriented towards the flashing light one of the neutral test passage would play from the same side of the booth. The sound file continued for approximately 12 seconds or until the infant looked away from the light for longer than 2 seconds. Shorter looks away were subtracted from the total looking time. All infants were tested

on the same neutral stimuli, thus any preference for one speaker over the other had to be due to the positive or negative affect the speakers expressed in the exposure phase.

Results and Discussion

In the test phase, 11 out of the 16 infants listened longer to the neutral voice of the formerly positive speaker than the neutral voice of the formerly negative speaker. A 2 X 2 repeated measures ANOVA was conducted to examine the effects of the speaker's emotional affect (positive or negative) and block (1st or 2nd block) on orientation time. The ANOVA revealed no main effect of emotional affect, $F(1, 15) = 2.47$, $p = .14$, $\eta_p^2 = .14$, or block $F(1, 15) = 1.19$, $p = .29$, $\eta_p^2 = .07$, however, there was a marginally significant interaction between mean orientation and block, $F(1, 15) = 3.80$, $p = .07$, $\eta_p^2 = .20$. Further analysis revealed this interaction was driven by differences in looking times to the speakers in the two blocks (see Figure 4). In Block 1, 12 out of the 16 infants oriented longer to hear the formerly positive voice. A two-tailed paired samples t-test indicated that the average looking times to the formerly positive speaker ($M = 9.16$, $SD = 4.04$) were significantly greater than the average looking times to the formerly negative speaker ($M = 6.95$, $SD = 3.12$), $t(15) = 2.52$, $p = .02$, $d = .63$. In Block 2, 8 out of the 16 infants oriented longer to hear the formerly positive voice. There was no significant preference for the formerly positive ($M = 7.35$, $SD = 4.04$) or the negative ($M = 7.22$, $SD = 3.39$) speaker, $t(15) = 0.14$, $p = .89$.

The preference for the neutral voice of the formerly positive speaker indicates that 10.5-month olds can use vocal affect information to selectively listen to a more socially desirable partner. The disappearance of this effect in Block 2 could be due to at least one of two reasons. It could be an indication that infants were fatigued, however, the experiment was relatively short and the dropout rate was low. It could also be that hearing the speaker's neutral voices repeatedly

in the test phase gradually dampened the effect of the emotional exposure in the first part of the experiment. This explanation raises the interesting possibility that infants might have perceived the tone of voice as a temporary emotional state rather than a stable personality trait of the speaker. As interesting as this last explanation is, we admit that it is speculative. No clear conclusion can be drawn from the current data, further experimentation would be needed to determine precisely why the effect disappeared in the second block of trials.

General Discussion

Social evaluation skills are highly developed early in life. At about the same time children take their first steps, or delight their parents with their first words, they are already keenly aware of the social environment that surrounds them. We have demonstrated that infants not only draw on visual information, but also use auditory information to form sensible judgments about others in their environment. Infants can actually retain information about a speaker's tone of voice and use it to guide their future social encounters with that individual. This finding stands in stark contrast to earlier work suggesting that even school-age children struggle to interpret vocal affect information. We suggest that the methodology used in the current study may simply be more effective at tapping into young children's vocal affect interpretation skills.

In Experiment 1, after being exposed to puppets that spoke in a positive and a negative affect, infants reached for the formerly positive puppet. In Experiment 2, when the puppets were replaced by inanimate sippy cups, infants no longer exhibited a preference. The results of Experiment 2 support our social interpretation of Experiment 1, suggesting that infants' puppet selections are evidence of social evaluation rather than simple association. These preferences for the formerly positive puppet and speaker go beyond studies showing infants prefer to listen to

positive over negative vocalizations (Papoušek et al., 1990) and are compatible with studies suggesting that infants will choose to interact with individuals that display positive 'pro-social' over 'anti-social' behavior (Hamlin et al., 2007; see also Kinzler et al., 2007, for related findings).

In Experiment 3, we tested infants in a paradigm involving exposure to voices only (i.e., not puppets). In this case, infants once again demonstrated that they can evaluate others based on vocal affect cues. After the exposure phase, infants responded differently to the neutral voices of the two speakers, orienting longer to the neutral voice of the formerly positive speaker than the neutral voices of the formerly negative person. It is interesting to note that in order to accomplish this task, infants must have recognized the speakers' voices in the test phase even though they were speaking in a different emotional affect. Although many studies have examined infants' ability to recognize words when produced in different emotional affects (e.g., Singh, Morgan, & White, 2004), this is the first study we know of to begin to examine infants' recognition of voices across different emotional affects. This is impressive given that even adults can struggle to recognize newly learned voices in certain contexts (Öhman, Eriksson, & Granhag, 2013; Philippon, Cherryman, Bull, & Vrij, 2007).

Taken together, the results of the three experiments presented in this paper demonstrate that at around the same time as infants are beginning to evaluate others based on visually observable behaviors, they are also learning to make similar judgments based on vocal affect information coming through the auditory channel. Although previous research has demonstrated that 4-month-olds prefer positive over negative vocalizations (Papoušek et al., 1990) this is the first study to show that older infants can actually retain this information and use it to guide their future social encounters with those individuals. It is possible that these early perceptual

preferences for positive over negative vocalizations may lay the foundation for the development of social preferences for individuals who speak in a positive tone of voice. A similar trajectory is seen for the development of other types of social preferences. For example, infants prefer to listen to their native language from birth (e.g., Moon, Cooper, & Fifer, 1993), fine-tune their voice and language recognition skills by 5 to 8 months (e.g., Johnson, Westrek, Nazzi, & Cutler, 2011; Nazzi, Jusczyk, & Johnson, 2000), and begin to show social preferences for individuals that speak in their native language by 6 to 10 months (Kinzler et al., 2007).

The study of infant social cognition is often limited to the use of behavioral measures (such as selection and looking time measures) to infer social understanding. Although, some have questioned the legitimacy of this approach, arguing that infants' reactions to these events could be driven by low-level attentional factors and simple associations rather than complex social evaluations (Haith, 1998; Scarf, Imuta, Colombo & Hayne, 2012), here we provide convergent evidence in support of higher-level social evaluation abilities by combining data from a social selection paradigm with looking time data. Moreover, the results we have reported in the current study align with several published studies using different methodologies and presenting stimuli in different domains (e.g. Hamlin et al., 2007).

Infant social cognition is a relatively new field of study. We are only beginning to explore the full extent of children's early social evaluation abilities. Future work should investigate how infants integrate and weight different social cues coming from different modalities and whether infants can retain this information about potential social partners over longer delays. It will also be important to further examine whether infants are responding differently depending on whether the emotion expressed is a stable trait of the character (the

character is an irritable person) or simply a brief emotional state (the character is irritated in response to a specific event).

Conclusions

The current study offers an important advance in our understanding of early social development. In three experiments we have demonstrated that social evaluation based on vocal affect emerges in infancy. This study goes beyond simple preferences for positive and negative affects to show that infants are able to use tone of voice information to guide their future social interactions with a speaker. Future research will seek to better understand the full scope and limitations of infants' social evaluation abilities, and further examine infants' reliance on visual versus auditory information in social judgment tasks. By better understanding normal social development, we hope to one day better understand abnormal patterns of development. Indeed, one can imagine that a variant of this paradigm could one day be part of a battery of tests used to screen infants at risk for developing autism.

References

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64,* 431-441. doi: 10.1037/0022-3514.64.3.431

Berman, J. M. J., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology*, *107*, 87-99. doi: 10.1016/j.jecp.2010.04.012

Bryant, G. A., & Barrett, H. C. (2007). Recognizing intentions in infant-directed speech: Evidence for universals. *Psychological Science, 18,* 746-751. doi: 10.1111/j.1467-9280.2007.01970.x.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146,* 668–670. doi: 10.1126/science.146.3644.668

Fawcett, C., & Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant and Child Development, 21,* 239-249. doi: 10.1002/icd.739.

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition, 110,* 160–170. doi: 10.1016/j.cognition.2008.11.010.

Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *NeuroReport*, *16*, 1825–1828. doi: 10.1097/01.wnr.0000185964.34336.b1

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development, 21*, 167-179. doi: 10.1016/S0163-6383(98)90001-7

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557-559. doi: 10.1038/nature06288

Hobson, R.P., Ouston, J., Lee, A. (1998). Emotion recognition in autism: coordinating faces and voices. *Psychological Medicine, 18,* 911-923. doi: 10.1017/S0033291700009843

Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14*, 1002-1011. doi:10.1111/j.1467-7687.2011.01052.x.

Johnson, E. K., & Zamuner, T. (2010). Using infant and toddler testing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.) *Experimental Methods in Language Acquisition Research* (pp. 73-94). Amsterdam: John Benjamins Publishing Company.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *The Proceedings of the National Academy of Sciences of the United States of America*, *104*, 12577-12580. doi: 10.1073/pnas.0705345104

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. Infant Behavior and Development, 16, 495–500. doi:10.1016/0163-6383(93)80007-u

Morton, J. B. & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development, 72*, 834–843. doi: 10.1111/1467-8624.00318

Morton, J. B., Trehub, S. E., & Zelazo, P. D. (2003) Sources of inflexibility in 6-year-olds' understanding of emotion in speech. *Child Development, 74*, 1857-1868. doi: 10.1046/j.1467-8624.2003.00642.x

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development, 67*, 3219-3237. doi: 10.1111/j.1467-8624.1996.tb01910.x

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43*, 1–19. doi:10.1006/jmla.2000.2698

Öhman, L., Eriksson, A., & Granhag, P. A. (2013). Angry voices from the past and present: Effects on adults' and children's earwitness memory. *Journal of Investigative Psychology and Offender Profiling, 10*, 57–70. doi:10.1002/jip.1381

Pascalis, O., & Kelly, D. J. (2009). The origins of face processing in humans phylogeny and ontogeny. *Perspectives on Psychological Science, 4*, 200–209. http://doi.org/10.1111/j.1745-6924.2009.01119.x

Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development, 13*, 539-545. doi: 10.1016/0163-6383(90)90022-Z

Philippon, A. C., Cherryman, J., Bull, R., & Vrij, A. (2007). Earwitness identification performance: The effect of language, target, deliberate strategies and indirect measures. *Applied Cognitive Psychology, 21,* 539–550. doi:10.1002/acp.1296

Premack, D., & Premack, A. J. (1997). Infants attribute value +/- to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience, 9,* 848–856. doi: 10.1162/jocn.1997.9.6.848

Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012). Social evaluation or simple association? Simple associations may explain moral reasoning in infants. *PLoS ONE, 7,* e42698. doi: 10.1371/journal.pone.0042698

Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3,* 365-394. doi: 10.1207/S15327078IN0303_5

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*, 173–189. doi:10.1016/j.jml.2004.04.004

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science, 7*, 261-269. doi: 10.1111/j.1467-7687.2004.00344

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development, 6*, 491-498. doi: 10.1016/S0163-6383(83)90331-4

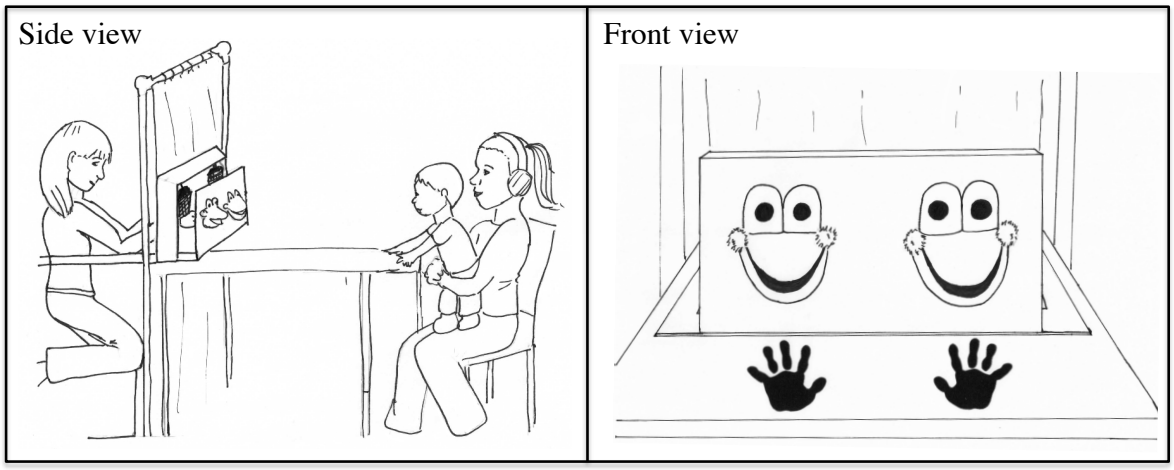


Figure 1. Experiment 1 diagram. Panel 1 depicts a side view of the experimental display. During the puppet exposure the experimenter (located behind the curtain) opened and closed the mouths of the puppets in synchrony with the auditory passages. The audio speakers were positioned directly behind the left and right puppets to give the illusion that the passages were coming from the puppets. The two puppets were closely matched in appearance but were different in color. Panel two illustrates the child's view of the puppet display prior to selection. The experimenter has moved the puppets within reach of the child using the wooden guides on the sides of the puppet display. Note that the puppets are equidistant from each other and from the hand print markings on the table.

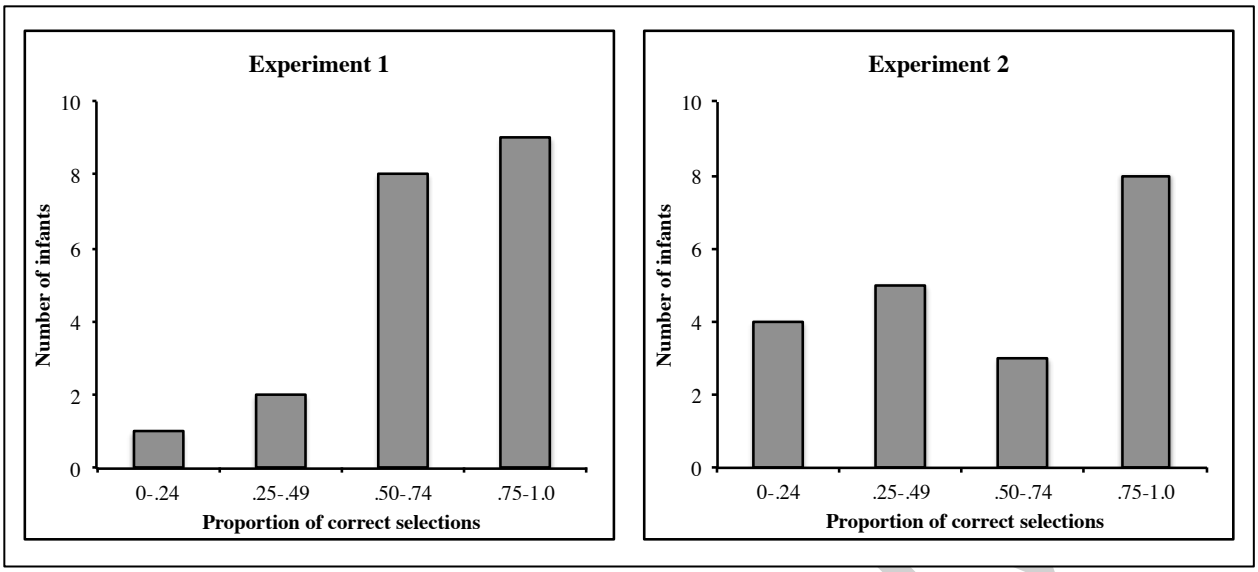


Figure 2. Frequency distribution of the proportion of correct selections in Experiments 1 and 2.

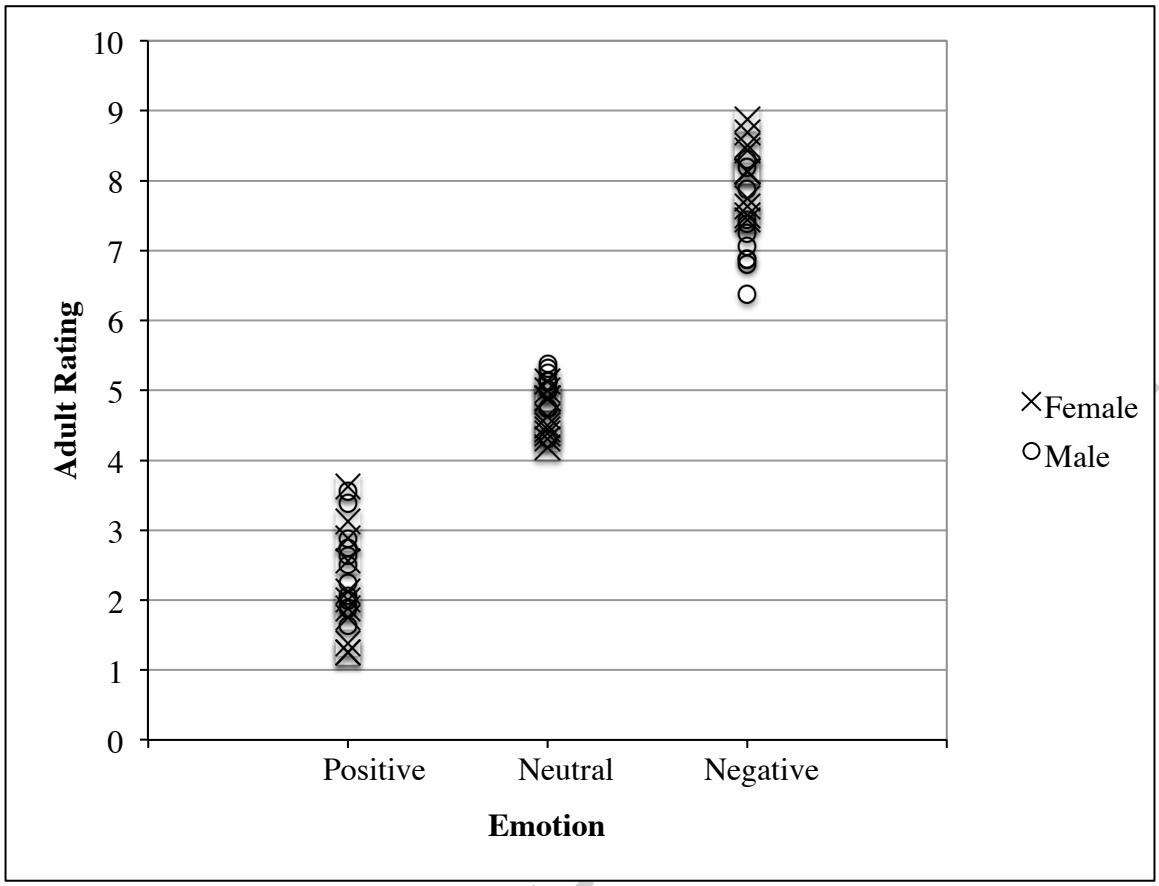
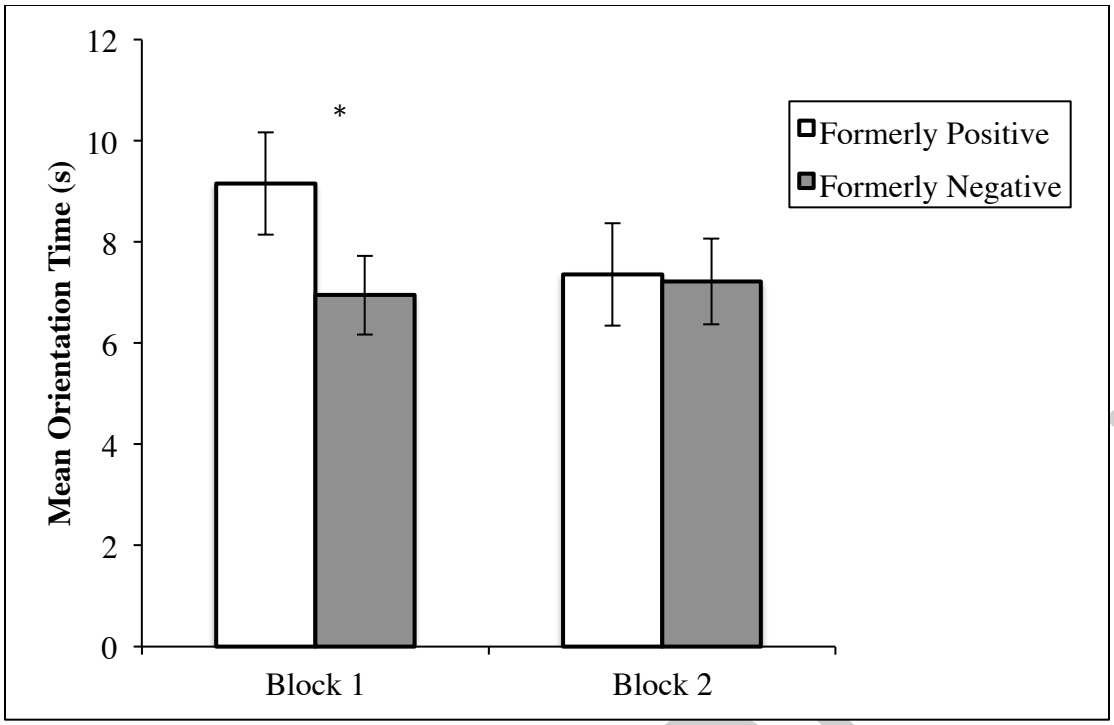


Figure 3. Average adult ratings of the positive, negative and neutral passages produced by both the male a female speaker. Eleven adults (M_{Age} = 19 yrs; 9 female, 2 male) rated each passage on a scale from 1(positive/happy) to 9 (negative/irritated) with 5 being neutral. Each point represents the average of one adult's ratings of 8 passages produced by one speaker in a given emotion.



* Two-tailed $p < .05$

Figure 4. Mean orientation time to the neutral voices of the formerly positive and formerly negative speakers in the Experiment 2. Since all infants were presented with the same neural test items, the observed preferences had to be a result of the affect (positive or negative) the speakers expressed in the exposure phase. Error bars represent the standard error of the mean.

Appendix

The 8 semantically neutral sentences that were recorded in happy, negative and neutral affects by both speakers.

1. You need to put on your jacket and mittens, it's cold outside.
2. Let me tie your shoe. You're going to trip and fall with it untied.
3. Go over and thank the nice lady for helping you find your coat.
4. You have cookie crumbs all over your face, let me wash them off.
5. We read this book last week, but we can read it again if you want to.
6. Put all your toys away, we need to tidy up before Dad gets home.
7. It's ok, we can come back and play at the park tomorrow.
8. If you sit really quietly now, then we can get ice cream after.