Child Development, September/October 2005, Volume 76, Number 5, Pages 1116–1126

Experiential Influences on Multimodal Perception of Emotion

Jessica E. Shackman and Seth D. Pollak
University of Wisconsin—Madison

The impact of 2 types of learning experiences on children's perception of multimodal emotion cues was examined. Children (aged 7–12 years) were presented with conflicting facial and vocal emotions. The effects of familiarity were tested by varying whether emotions were presented by familiar or unfamiliar adults. The salience of particular emotional expressions was tested by contrasting the performance of physically abused and nonabused children. Children exhibited a preference for auditory expressions produced by their mothers but not by strangers. Additionally, abused children were biased to rely on auditory cues when their own abusive mother was expressing anger. These results are discussed in terms of the impact of both typical and atypical early experiences on the development of emotion perception.

Children are confronted early in life with the task of learning to decode and make sense of multiple simultaneously presented emotional signals. Emotional information is conveyed in the form of linguistic (i.e., semantic content of spoken language) and paralinguistic (e.g., facial expression, vocal prosody, physical gestures, and body posture) cues. These affective signals are typically congruent, and this redundancy facilitates efficient processing of multiple emotional signals (Bahrick & Lickliter, 2000; deGelder, Böcker, Tuomainen, Hensen, & Vroomen, 1999). Indeed, infants are able to accurately perceive multimodal signals in the first months of life (e.g., Lewkowicz, 1996). Moreover, although inputs from different modalities are processed in separate areas of the brain, our conscious experience is one of coherent, unified perceptions, reflecting that information becomes integrated across sensory modalities (Fingelkurts, Fingelkurts, Krause, Moettoenen, & Sams, 2003). Indeed, auditory and visual information are efficiently integrated when sensory information is redundant (Bahrick, Lickliter, & Flom, 2004). However, when sensory channels convey distinct or conflicting information, perceivers must resolve conflict between signals by preferentially attending to one source of information. Little is currently known about the role of the developing child's experiences in learning to respond to multiple emotion signals.

Most research on the development of emotion perception has focused on children's recognition of facial expressions. Faces are certainly important in the communication of emotion. Darwin (1872/2002) argued that facial expressions originate in basic acts of self-preservation common to human and nonhuman animals; indeed, humans have particularly well-developed abilities to accurately recognize emotional facial expressions (Ekman, 1993). Yet, facial expressions do not occur in isolation—humans must routinely rely on multiple sources of information during social communication (Massaro & Egan, 1996). In addition to faces, the acoustic properties of speech make the voice a rich source of information about an individual's affective state. Vocal expressions of emotions are particularly important from a developmental perspective in that auditory signals can capture attention from someone who is not already visually attending to the expresser, as is often the case in the communications between infants and toddlers and their caretakers. Despite the salience of auditory cues, Fernald (1993) has argued that little attention has been directed towards understanding the role of vocal expressions in emotion perception.

Listeners are able to accurately recognize joy, sadness, anger, and fear based upon prosody, which includes both the rhythmic and intonational aspects of human speech, when listening to semantically

The authors wish to thank Jenny Saffran for helpful comments on an earlier version of this manuscript. We also appreciate the assistance of Anna Bechner and Andrew Goodman in the collection of these data, and gratefully acknowledge Martha Alibali and Alexander Shackman for discussions regarding the analysis of these data. These experiments would not be possible without the participation of many children and their families, for whose collaboration we are extremely appreciative. This research was supported by grants from the National Institute of Mental Health to Seth Pollak (MH61285, MH68858). Jessica Shackman was also supported by a National Institutes of Mental Health Training Program in Emotion Research (MH18931), and a University of Wisconsin Graduate Fellowship.

Correspondence concerning this article should be addressed to Seth D. Pollak, 1202 West Johnson Street, University of Wisconsin at Madison, Madison, WI 53706-8190. Electronic mail may be sent to spollak@wisc.edu.

neutral sentences (Johnson, Emde, Scherer, & Klinnert, 1986). Although listeners tend to associate particular acoustic cues with discrete emotions, such vocal patterns in isolation are not always reliable indices of emotion (Scherer, Banse, Wallbott, & Goldbeck, 1991). In light of this, Bachorowski (1999) has suggested that the talker–listener relationship and the speaker's intended impact of vocal signals on the affective state of the listener are also important considerations. Thus, vocal and facial expressions together can provide information above and beyond what can be gleaned from either modality alone.

Multimodal Affect Perception and Development

The importance of the reciprocal influence of facial movements and vocal expressions was first demonstrated with the McGurk effect in the area of speech perception (McGurk & MacDonald, 1976). This perceptual effect is obtained when a subject is simultaneously presented with a visually articulated syllable that differs from the one presented acoustically. In such situations, observers combine multiple sources of information into an integrated, coherent percept not present in either modality alone. The McGurk effect has also been extended and used as a model of bimodal emotion perception. Studies on this topic suggest that in adults, facial expressions tend to be more influential than vocal expressions (e.g., Bugental, Kaswan, Love, & Fox, 1970; Hess, Kappas, & Scherer, 1988), yet the effectiveness of a given modality has been shown to vary to the extent that the other modality is ambiguous (Massaro & Egan, 1996) and incongruent (deGelder & Vroomen, 2000). It has been proposed that bimodal perception occurs in three stages: evaluation, integration, and decision making (the Fuzzy Logical Model of Perception, FLMP; Massaro & Egan, 1996). First, each separable source of information is evaluated based on prototypes of particular emotional expressions. Next, integration involves the combination of the degree to which each source supports a given alternative (e.g., happy, angry, sad). Finally, a decision is made based on the amount of support for each alternative. When one source of information only weakly supports a possible alternative, other sources of information are given more influence. Thus, studies with adults have provided us with a mechanistic model of the circumstances under which visual versus auditory information is more influential.

Although adults are able to flexibly attend to both auditory and visual information, most individuals show perceptual and attentional biases towards visual stimuli (Robinson & Sloutsky, 2004). But these attentional and perceptual processes undergo developmental changes (Barrett, Campos, & Emde, 1996; Kahana-Kalman & Walker-Andrews, 2001; Soken & Pick, 1992). In contrast to adults, infants as young as 6 months of age are more influenced by auditory than visual input when discriminating emotional expressions (Caron, Caron, & MacLean, 1988; Lewkowicz, 1988). Similarly, 5-month-old infants respond to affective vocal expressions presented in the absence of facial expressions, but not vice versa (Fernald, 1993). These behavioral findings are consistent with the earlier development of the auditory, as compared to the visual, system (DeCasper & Fifer, 1980). Emotion perception through facial channels may lag behind auditory channels because the visual system cannot detect fine detail until 3 months of age, and does not develop sufficient contrast sensitivity to detect static facial expression contrasts until almost 6 months (Hainline & Abramov, 1992; Walker-Andrews, 1997).

Taken together, extant research suggests that infants are more influenced by auditory information and adults are more influenced by visual information. But affect perception in middle childhood is poorly understood, leaving the developmental picture of these processes incomplete. Studies of exposure to media indicated that children tended to recall more visual than auditory information (Hayes & Kelly, 1984; Hayes, Kelly, & Mandel, 1986; Pezdek & Stevens, 1984). But these studies evaluated recall, rather than perceptual processing, of information. Perceptual studies suggest that when presented with nonemotional stimuli, 5-year-old children rely predominantly on auditory cues relative to visual information (Sloutsky & Napolitano, 2003). However, modality dominance shifts flexibly in young children, depending on the stimulus conditions. For example, 4-year-old children are likely to process stimuli only in the preferred modality when both auditory and visual information are of equal salience, suggesting that auditory and visual information compete for attentional resources early in development (Robinson & Sloutsky, 2004).

The Emotion Learning Environment and the Development of Perceptual Abilities

One way to evaluate how emotion perception abilities are shaped by children's experiences is to examine the effects of children's familiarity with the person conveying the emotion. Children's earliest exposure to emotional expressions often occurs in the context of familiar family members, as infants observe and learn

to interpret and respond to the expressive behaviors of their caregivers (Fox, 1994; Montague & Walker-Andrews, 2002). By 1 year of age, infants are able to use facial cues produced by their caregivers to evaluate potential threat, as evidenced by ''social referencing'' behaviors (e.g., Klinnert, Emde, Butterfield, & Campos, 1986; Sorce, Emde, Campos, & Klinnert, 1985). Indeed, familiarity with the individual expressing an emotion enhances the infant's ability to extract meaningful information from multimodal emotional expressions. For example, $3\frac{1}{2}$-month-old infants were better able to recognize facial expressions accompanied by affectively matched vocal expressions, but only when the expressions were produced by the infants' mothers (Kahana-Kalman & Walker-Andrews, 2001). Similarly, the ability of infants to accurately pair facial expressions of emotion with affectively concordant or discordant vocal expressions was correlated with the amount of parent–infant contact time (Montague & Walker-Andrews, 2002).

The role of familiarity in perception is not limited to infancy. Adult listeners make more detailed and accurate perceptual judgments of emotion in the voice for familiar than for unfamiliar talkers (Bachorowski, 1999), familiar stimuli are more likely than novel stimuli to engage attention (Napolitano & Sloutsky, 2004), and familiar stimuli elicit a neuronal response in the primate brain that is almost twice the magnitude of the response to novel stimuli (Höscher, Rolls, & Xiang, 2003). However, the relative importance of familiarity in emotion perception may also change with development. For example, infants exhibit a larger Nc component of the event related potential (believed to index increased attention) to their mother's face relative to a stranger's face, whereas preschool-age children show a larger Nc to unfamiliar faces (Carver et al., 2003). Thus, infants appear to be sensitive to contextual information that facilitates extraction of meaning from emotional expressions posed by familiar individuals. This may reflect the fact that maternal expressions of emotion are more informative with respect to ensuing actions. But the motivational significance, and processing of this information, may change with development.

A second way to evaluate the role of experience is to index the effect of children's familiarity with particular emotions. Although it is difficult to measure the amount of exposure any individual has had to particular emotions with precision, it is possible to estimate gross differences across groups. For example, studies of maltreating mothers suggest that the emotional and expressive environments they provide for their children deviate in important ways from normal experience (Camras, Sachs-Alter & Ribordy, 1996). Abusive parents engage in fewer positive emotional interactions with their children than nonabusive parents (Burgess & Conger, 1978) and direct more negative affect toward their children than nonmaltreating mothers (Trickett, Aber, Carlson, & Cicchetti, 1991). Although maltreated children are generally poorer at recognizing emotions (Camras et al., 1990), the experience of physical abuse may increase children's attention to anger (Pollak, Klorman, Thatcher, & Cicchetti, 2001).

Hypotheses

The aim of the present experiment is to examine learning factors that may influence the perception of multimodal emotional expressions among school-age children. To this end, we evaluated two experience-based factors that might influence how children perceive auditory and visual emotional information: learning about how particular individuals express emotions and learning about the salience of particular emotional expressions. Based upon extant literature, which suggests that children's modality preferences are contingent upon characteristics of the stimuli presented, we expected that children's perceptual behavior would interact with both familiarity of the expresser as well as the salience of the emotion presented.

The influences of familiarity with how particular individuals express emotion were tested by varying whether emotions were presented by familiar or unfamiliar adults. Based upon findings that the talker–listener relationship is especially important in the recognition of vocal affect (e.g., Bachorowski, 1999), we predicted that children would rely more on auditory signals relative to facial signals when emotions were expressed by familiar adults. In contrast, because children may be able to extract less meaningful information from an unfamiliar auditory expression and have been shown to attend more to unfamiliar faces (e.g., Carver et al., 2003), children were expected to rely more on facial cues when decoding emotions expressed by a stranger.

The salience of particular emotions was evaluated by contrasting the performance of children who are developing within unusually hostile family environments to peers developing within more normative contexts. Such an analysis also addresses mechanisms underlying the development of social information processing skills. Anger is likely to acquire enhanced salience in an abusive environment. Because abused children exhibit perceptual sensitivity for visual displays of anger (Pollak & Sinha,

2002), we predicted that physically abused children would show a perceptual preference for anger over other emotional expressions. We reasoned that children might be especially senstive to these expressions when they were produced by their abusive mothers because this threat cue has been associated with salient physical consequences. We expected that the interaction between group, emotion, and familiarity would apply to both vocal and facial expressions of anger.

Method

Participants

Sixty-three mothers and their children ranging in age from 7 to 12 years ($M = 9.57$, $SD = 1.83$) participated in this study. The physically abused group included 33 children with substantiated cases of child maltreatment recruited through a county Child Protective Services agency. Mothers were the verified abusive parent in all of these cases. These children were compared with 30 nonabused children. Efforts were made to recruit 2 samples with similar family demographics (see Table 1). All children had normal or corrected-to-normal vision. Just prior to testing, each child's hearing was screened following guidelines of the American Speech-Language-Hearing Association (ASHA).

Nonabusive mothers were screened through the registry of state Child Protective Services records and with the Parent–Child Conflict Tactics Scale (PCCTS; Straus, Hamby, Finkelhor, Moore, & Runyan, 1998), a measure of the extent to which a parent has carried out specific acts of physical and verbal aggression toward the child. Based on the PCCTS, parents of children in the control group did not endorse aggressive behaviors toward their children ($M = 2.33$, $SD = 2.88$); the opposite was true of the physically abusive parents ($M = 34.97$, $SD = 23.12$), $t(61) = 7.67$, $p < .01$. The aggressive behaviors endorsed by parents of children in the control group were limited to items reflecting corporal punishment, including spanking.

Table 1
Means (and Standard Deviations) of Sociodemographic Characteristics of Sample

Characteristic	Physically abused	Control
Number of boys/girls	17/16	15/15
Age (years)	9.73 (1.7)	9.39 (1.9)
Socioeconomic status[a]	32.5 (10.1)	38.2 (6.1)
Race (%)		
African American	60	30
Caucasian	40	70

[a]The Four Factor Index of Social Status (Hollingshed, 1975) reflects family socioeconomic status on the basis of parent education and occupational status.

Parents received detailed information about the study protocol before giving informed consent. After being shown the study apparatus, children gave verbal assent for participation. Children were rewarded with age appropriate prizes, and families received $20 for their participation in the study.

Stimuli

Facial and vocal stimuli were created individually for each child. Each child's mother visited the laboratory prior to her child's participation in the study for the purpose of stimulus development. Mothers were asked to recall a memory or imagine a situation that would help them to accurately express the desired emotion on their face and in their voice, and were given the opportunity to practice their emotional expressions using a mirror. Facial images were recorded using a Sony Mavica digital camera (MVC-CD400) while mothers posed happy, sad, and angry facial expressions. Mothers then reviewed the photographs and, along with the experimenter, selected the images they felt best depicted their targeted emotional state. Digital photographs were then edited with Adobe Photoshop so that each facial stimulus was similar in size, contrast, and luminance. Mothers next read semantically neutral sentences with happy, sad, and angry tones of voice. Vocal stimuli were recorded with a Sony MiniDisc recorder (MZ-N1) and edited using Cool-Edit software to equalize the volume and length of each utterance.

To examine whether potential differences in mothers' posing abilities might account for differences in children's task performance, 100 undergraduate students (68 females) rated each of the facial and vocal expressions produced by the mothers in this study. Subjects were asked to rate each expression, based on its prototypicality, on a 10-point scale. Before rating the expressions, raters were presented briefly with all of the stimulus items. Raters viewed two presentations of each face and then heard each of the five vocal samples, grouped by emotion.

Procedure

Children were tested individually during the afternoon, in a dimly lit sound attenuated room, at a distance of 1 m from the computer screen, positioned so that the stimuli appeared on the subject's horizontal straight-ahead line of sight. Facial images

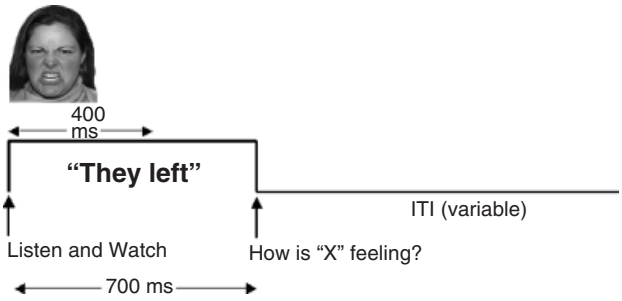

Figure 1. Schematic of experiment structure and timing.
Note: The face in this figure is not the photograph of any participating mother, nor was it used at any point during the experiment. It is presented in this figure solely for the purpose of illustrating the experiment structure.

were presented on a 19-inch Viewsonic monitor and occluded a space equivalent to 500 × 600 pixels, with the screen resolution set to 1024 × 768 pixels. Auditory stimuli were presented through Optimus LV-20 earphones. On any given trial, children saw (a) their mother's face expressing happiness, anger, or sadness paired with either the same or a different vocal expression or (b) the face of another mother, who was unfamiliar to the participant, expressing happiness, anger, or sadness paired with the same or a different vocal expression. Familiar faces were always paired with familiar voices, and unfamiliar faces were always paired with unfamiliar voices. All children saw and heard their own mother as well as one unfamiliar nonmaltreating mother. Unfamiliar stimuli were mothers of other children who participated in this experiment, matched for race.

Trials began with the simultaneous onset of a voice, lasting for approximately 700 ms, and a face, presented for 400 ms (Figure 1). Each of the 9 stimulus combinations (3 vocal emotions × 3 facial emotions × 2 mothers) was presented 5 times in pseudorandom order across two 45-trial blocks. The intertrial interval between voice offset and the onset of the next face and voice was based upon each child's response time for each trial, and was manually controlled by the experimenter. Following the experimental trials, children completed a control task consisting of 2 blocks of 30 trials each. In these trials, each facial and vocal expression (for mother and stranger) was presented alone to ensure children could identify the emotions conveyed by the stimuli. Children were asked to verbally report what each person was feeling. Instructions were worded so as not to emphasize one modality over the other, and to not constrain the emotion labels children might provide. Children's verbal labels of "angry," "mad," and "irritated" were classified as angry; "happy," "excited," and "glad" were classified as happy; and the responses "sad," "depressed," "unhappy," and "down" were classified as sad. Other responses were rare, considered incorrect, and did not factor into children's scores.

Results

The aim of the current experiment was to examine how two types of affective learning experiences, familiarity with the emotion poser and familiarity with particular emotions, affect the development of children's perception of multimodal affective information. We first present results examining modality preferences in children. Then, to evaluate the role of learning factors, we present data on how familiarity with the way in which particular people express emotions and familiarity with particular emotional expressions influence these modality preferences. Data on modality dominance are referred to as perceptual preferences, a term that reflects children's increased sensitivity or attentiveness to particular stimuli, leading to preferential processing of one stimulus over another. Finally, to address additional factors in children's perceptual processing, we present data on children's perception of unimodal information and preliminary data on stimulus characteristics and the relationship to children's responses.

Perceptual preference scores calculated for each child reflected the frequency with which children correctly identified vocal and facial emotions. To generate these scores, all trials containing a particular emotion in a given modality were combined. For example, out of all of the trials containing happiness in the face, scores were computed by dividing the total number of times children identified the face–voice pair as "happy" by the total number of trials. Therefore scores ranged from 0 to 1.0, with a score above 0.5 indicating a preference for happiness in the face. A similar analysis was conducted to evaluate children's responses to happiness in the voice (and other emotions). All subjects responded to congruent pairs with 100% accuracy; therefore, these trials were not subjected to statistical analyses; these trials ensured that all subjects could maintain attention throughout the experiment (significant decrements in accuracy would have indicated lapses in attention). Responses in which a child identified an emotion that was not present in either the face or the voice (e.g., if a child responded "afraid" or "surprised" to an angry voice paired with a sad face) were also recorded, but were rare and considered incorrect. The mean error rate across all children was 4.5% ($SD = 4.6\%$), and the two groups did not differ significantly on their error rates, $t(61) = .01$, *ns*.

Scores were submitted to an analysis of variance with maltreatment group (control, physical abuse) as a between-subjects factor, and with emotion (angry, happy, sad), person-familiarity (mother, stranger), and modality (auditory, visual) as within-subjects factors. Probability values for repeated measures are reported with Greenhouse–Geisser corrections. The abuse and control groups differed in terms of racial composition, $t(61) = 2.53$, $p < .05$, and socioeconomic status, $t(61) = 2.73$, $p < .05$. Therefore, these factors were included in initial analyses. Child race, gender, and SES were not related to any outcome variable (all $ps > .06$), and for simplicity, are not included in the analyses reported here. The results reported below remained unchanged when race and socioeconomic status were included in the ANOVA.

Modality Preferences

In general, children did not show a preference for either auditory or visual information, $F(1, 61) < 1$. However, children's modality preferences were moderated by the emotions being expressed, $F(2, 122) = 33.80$, $p < .001$. Children identified happy facial expressions ($M = .57$, $SD = .04$) more frequently than happy vocal expressions ($M = .39$, $SD = .02$), $F(1, 61) = 9.25$, $p < .01$. However, no modality preferences were evident for either anger, $F(1, 61) < 1$, or sadness, $F(1, 61) < 1$.

Effects of Poser-Familiarity

Children's modality preferences were moderated by their familiarity with the person posing the emotion, $F(1, 61) = 33.69$, $p < .001$. As shown in Figure 2, children preferentially processed the face when the person expressing the emotion was unfamiliar, $F(1, 61) = 29.82$, $p < .001$, and preferentially processed the voice when the person expressing the emotion was familiar, $F(1, 61) = 35.13$, $p < .001$. An interaction between emotion and familiarity also emerged, $F(2, 122) = 13.81$, $p < .001$, reflecting that children identified expressions of anger ($M = .52$, $SD = .02$) and happiness ($M = .51$, $SD = .02$) posed by their mothers more frequently than those posed by a stranger (anger, $M = .44$, $SD = .02$; happiness, $M = .45$, $SD = .02$), across both modalities. However, sadness was perceived more frequently when posed by an unfamiliar adult (mother, $M = .38$, $SD = .02$; stranger, $M = .50$, $SD = .02$).

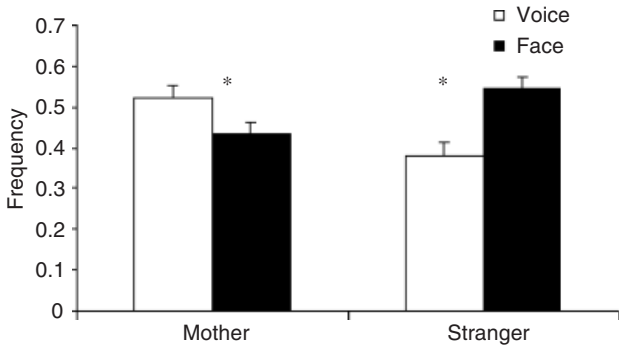

Figure 2. Modality preferences (reflected in frequency of identification) for familiar and unfamiliar expressions of emotion. A frequency score of 0.5 indicates that children showed an equal preference for both auditory and visual information. * Indicates $p < .05$.

Effects of Emotion Salience

We hypothesized that abused children would show preferential processing of anger, and this was confirmed by an interaction between group and emotion, $F(2, 122) = 5.49$, $p < .01$. Planned comparisons revealed that abused children preferentially processed anger more than controls (Abuse $M = .53$, $SD = .02$; Control $M = .43$, $SD = .02$), $F(1, 61) = 10.33$, $p < .01$. The two groups did not differ in their identification of happiness or sadness. However, this effect was further qualified by a three-way interaction between maltreatment group, emotion, and familiarity, $F(2, 122) = 4.26$, $p < .02$ (see Figure 3). In particular, abused children identified anger expressions more frequently when produced by their own mothers, $F(1, 32) = 18.68$, $p < .001$. This effect did not emerge for controls, nor did it apply to abused children's identification of happiness, Fs $(1, 29) < 1$; however, abused children identified sadness less frequently in their own mothers than in an unfamiliar adult, $F(1, 32) = 36.3$, $p < .001$. Additionally, abused children identified anger expressions more frequently than did control children when expressions were produced by their own mothers, $F(1, 61) = 16.27$, $p < .001$, but not when expressions were produced by unfamiliar women, $F(1, 61) < 1$.

To test our hypothesis concerning group differences in the perception of anger, we next examined children's emotion identification only on trials where anger was present. This analysis revealed an interaction among group, modality, and familiarity on trials containing only anger, $F(1, 61) = 3.53$, $p = .06$. As shown in Figure 4, physically abused children identified their mothers' vocal anger more frequently than vocal anger expressed by a stranger, $t(32) = 6.70$, $p < .001$, but no differences were observed for abused children's recognition of facial anger, $t(32) = .11$, *ns*. Likewise, abused children

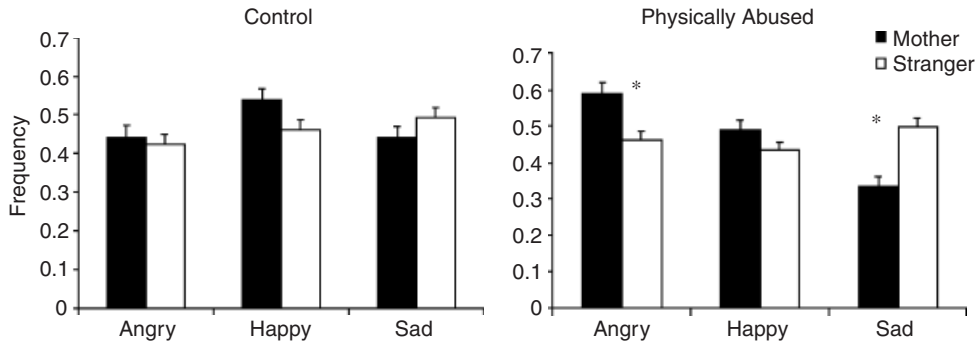

Figure 3. Depicts group × emotion × familiarity interaction. Physically abused children (right) made significantly more anger identifications for expressions produced by their own mothers, relative to unfamiliar expressions, $F(1, 32) = 18.68$, $p < .001$. No differences emerged for control children's (left) identifications of familiar and unfamiliar expressions of anger, $F(1, 29) < 1$. Additionally, the group difference for familiar expressions of anger was also significant, $F(1, 61) = 7.69$, $p < .01$.* Indicates $p < .01$.

identified their mothers' vocal anger more frequently than their mothers' facial anger, $t(32) = 3.65$, $p = .001$, but no modality differences were observed for abused children's recognition of anger expressed by unfamiliar adults, $t(32) = .19$, *ns*. Frequency of anger identification among controls was similar for facial and vocal expressions, $F(1, 29) = 2.38$, *ns*. No significant interactions emerged among group, modality, and familiarity for trials containing happiness or sadness, Fs$(1, 61) < 1$; thus, this effect appears to be specific to anger.

Baseline Emotion Identification

To ensure that children in both groups could accurately identify each emotion, faces and voices were presented independently. As described above, children's scores were computed as a ratio of the number of correct identifications to the total number of trials containing each emotion presented. These scores were submitted to a repeated measures analysis of variance with maltreatment group (control, physical abuse), gender, and race as between-subjects factors, and emotion (angry, happy, sad), person-familiarity (mother, stranger), and modality (auditory, visual) as within-subjects factors. Child race, gender, and socioeconomic status were not related to recognition accuracy, all ps $> .09$. All subjects correctly identified facial and vocal stimuli with greater than 80% accuracy (Table 2). No significant accuracy differences emerged based on maltreatment group, $F(1, 55) = .09$, *ns*, emotion, $F(2, 110) = 1.53$, *ns*, modality, $F(1, 55) = .49$, *ns*, or familiarity, $F(1, 55) = .11$, *ns*, nor were any interactions observed between these variables, all ps $> .1$. Therefore, the results reported above are not likely to be confounded by differences in children's emotion recognition abilities.

Ratings of facial and vocal stimuli

We undertook a set of post hoc analyses to examine whether differences in mothers' posing abilities could account for differences in children's task

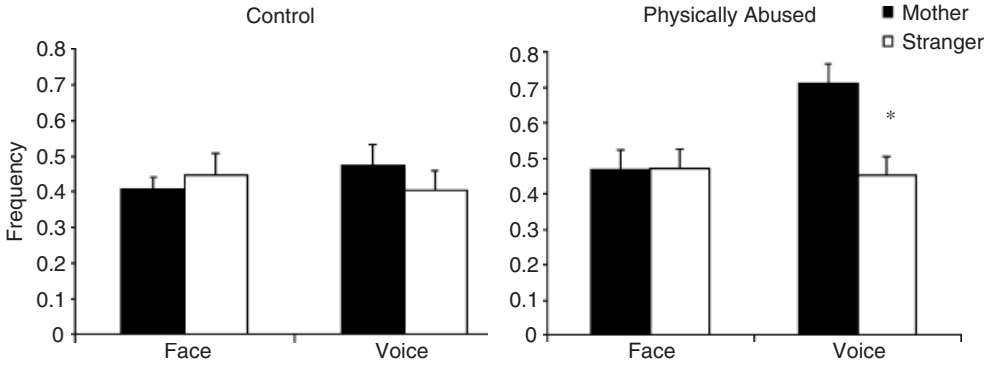

Figure 4. Children's modality preferences for *anger* expressions produced by their own mother and a stranger. Physically abused children identified significantly more anger expressions in their own mother's voice relative to unfamiliar vocal anger expressions, and relative to their own mother's facial expressions of anger. Control children did not show any modality preferences for anger identification. * Indicates $p < .01$.

Table 2
Children's Accuracy in the Unimodal Emotion Recognition Task; Values are Reported as Percentage Correct

	Facial emotion						Vocal emotion					
	Familiar			Unfamiliar			Familiar			Unfamiliar		
Group	Angry	Happy	Sad	Angry	Happy	Sad	Angry	Happy	Sad	Angry	Happy	Sad
Control	1.00	.99	.98	.99	.99	.98	.98	.99	.98	.99	.99	.98
Abused	.99	.98	.99	.99	.98	.99	.99	.98	.99	.98	.99	.97

performance. A repeated measures ANOVA was conducted on the rating scores provided by undergraduates for each stimulus item, with status of mother (abusive, nonabusive), modality (vocal, facial), and emotion (angry, happy, sad) as within-subjects factors. A three-way interaction between emotion, modality, and group, $F(2, 196) = 3.84$, $p < .03$, was further explored using paired-samples t tests to examine group differences in posing ability for each emotion expressed in each modality. Nonabusive mothers were rated as having more prototypical expressions of emotion than maltreating mothers for angry facial expressions, $t(99) = 2.18$, $p < .05$, and angry, $t(99) = 9.22$, $p < .001$, happy, $t(99) = 18.24$, $p < .001$, and sad vocal expressions, $t(99) = 6.92$, $p < .001$. We then computed correlations between the undergraduate ratings of each mother's posing ability for each emotion and modality, and the frequency with which their child identified that emotion. This analysis revealed that undergraduate ratings of mothers' posing abilities were unrelated to children's task performance: angry face, $r(62) = -.041$, *ns*, angry voice, $r(62) = .004$, *ns*, happy face, $r(62) = .219$, *ns*, happy voice, $r(62) = .202$, *ns*, sad face, $r(62) = .109$, *ns*, and sad voice, $r(62) = .110$, *ns*

Discussion

The present study investigated the effects of learning on children's perception of emotion. We approximated variations in the effects of children's experiences by examining how children responded to emotional signals expressed by familiar versus unfamiliar adults and the effects of salience by testing children living in abusive family environments. There was no evidence of general modality dominance in school-age children; that is, children did not exhibit consistent preferences for either visual or auditory information. The finding that school-age children do not display consistent modality dominance is in agreement with prior reports that modality preferences change with development (Robinson & Sloutsky, 2004) and raises the possibility that environmental factors may influence modality preferences during middle childhood. Rather, children exhibited an auditory preference when presented with emotions expressed by their mothers, and a visual preference for emotions expressed by a stranger. This is consistent with the idea that vocal emotion may be more difficult to identify when expressed by unfamiliar individuals (Bachorowski, 1999). Perceptual processing was also influenced by emotion: children demonstrated a preference for visual over auditory expressions of happiness. This finding is consistent with reports that although happiness is an easily recognizable facial expression (Ekman, 1994), it is more difficult to identify in the voice (Scherer et al., 1991; Scherer, Banse, & Wallbott, 2001). Finally, maltreated children demonstrated a bias toward a reliance on auditory cues when their abusive mother was expressing vocal anger. This finding is consistent with the hypothesis that modality dominance (and auditory dominance in particular) reflects automatic attention that is enhanced by salient information (deGelder & Vroomen, 2000; Napolitano & Sloutsky, 2004). Based upon findings from other studies of bimodal perception, it is possible that the meaning of an angry stimulus is less ambiguous for physically abused than for typically developing children. Therefore, anger may have more perceptual influence relative to other sources of social information among abused children.

The finding that abused children showed a preference for auditory anger only when these expressions were familiar supports the notion that contextual cues, such as familiarity, serve to enhance children's ability to extract meaningful information from emotional expressions. As noted in studies with infants (Kahana-Kalman & Walker-Andrews, 2001), maternal smiles are usually accompanied by positive caretaking behaviors. Likewise, abused children may be more motivated to attend to their mothers' vocal anger expressions because they may

typically foreshadow more specific and significant punishment outcomes for them. Through learning, this response may become automatic. The fact that this effect was not observed for familiar facial expressions of anger is worth noting, and we speculate on this below.

Prior research on the effects of atypical experience on emotional development has focused almost exclusively on facial expressions of emotion, and thus we know little about the relative influence of auditory emotional expressions on children's perceptual and attentional abilities. Although angry faces are a salient stimulus for physically abused children, it is important to consider the possibility that angry voices may be even more frequently experienced and associated with a wider variety of negative consequences. Typically, children encounter an angry face in the context of a direct interaction, when a threat is already present and anger is being directed at the child. In contrast, vocal expressions of anger may be experienced directly by the child, or indirectly in the context of a caregiver fighting or yelling at siblings. Angry voices may also serve as a warning signal that threat is imminent, since an angry voice can be detected from a greater distance than the distance from which one can recognize an angry face. Thus, angry voices may be a more reliable predictor of environmental change and negative consequences, such that they would allow a child to detect angry encounters sooner and more often, and to engage appropriate resources more efficiently in order to avoid threat. That physically abused children tended to be especially sensitive to vocal anger expressed by their abusive mother suggests some specificity in children's perceptual learning, such that those more familiar and meaningful stimuli may take precedence for perceptual processing over those stimuli that are less meaningful. Thus, research that employs standardized expressions of facial emotion may not capture some of the processes linking emotional experience to perception and behavior.

Consistent with prior reports (e.g., Camras et al., 1988), the present study suggests potentially important differences in the communication of emotion by physically abusive and nonmaltreating mothers. Physically abused children demonstrated an *enhanced* perceptual preference for expressions of anger, leading one to expect that maltreating mothers would be *better* at anger production. However, the current sample of physically abusive mothers was *less* able to produce good exemplars of anger (both facial and vocal), despite the fact that their children had no difficulty recognizing those same stimuli presented to each modality in isolation. Although this may at first seem like a contradiction, the fact that maltreating mothers produced poorer quality emotional signals may actually enhance the need for their children to fine tune their attention to more subtle signals of anger expressed by their mothers. This would then result in maltreated children's enhanced sensitivity to subtle changes in their mothers' voices, changes that may not be readily apparent to a nonabused child. In the light of this apparent contradiction, future efforts should continue to take a more fine-grained approach to understanding the differences in parents' expressive abilities to better understand the affective learning environments they provide for their children.

The research reported here suggests that the importance of auditory versus visual percepts is influenced by the meaning attached to particular emotions contained in each expression and by the familiarity of the individual expressing emotion. However, future research should examine whether children are able to deliberately control the deployment of their attention to auditory versus visual stimuli of unmatched salience or whether these shifts in attention truly occur automatically. Prior research has demonstrated that attention allocated toward anger in the environment appears to be critical to understanding the information processing deficits associated with child maltreatment. However, one limitation of prior work is the focus on facial expression of affect, to the exclusion of other emotional signals that children encounter in their environment. The investigation of the myriad ways in which children learn to perceive and attend to emotions will likely provide a more complete picture of the complex interactions between developing children and their environment.

References

Bachorowski, J. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science*, *8*, 53–57.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*, 190–201.

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.

Barrett, K. C., Campos, J. J., & Emde, R. N. (1996). Infants' use of conflicting emotion signals. *Cognition and Emotion*, *10*, 113–135.

Bugental, D. E., Kaswan, J. W., Love, L. R., & Fox, M. N. (1970). Child versus adult perception of evaluative

messages in verbal, vocal, and visual channels. *Developmental Psychology, 2*, 367–375.
Burgess, R., & Conger, R. (1978). Family interaction in abusive, neglectful and normal families. *Child Development*, *49*, 1163–1173.
Camras, L. A., Ribordy, S., Hill, J., Martino, S., Sachs, V., Spaccarelli, S., & Stefani, R. (1990). Maternal facial behavior and the recognition and production of emotional expression by maltreated and nonmaltreated children. *Developmental Psychology, 26*, 304–312.
Camras, L. A., Ribordy, S., Hill, J., Martino, S., Spaccarelli, S., & Stefani, R. (1988). Recognition and posing of emotional expressions by abused children and their mothers. *Developmental Psychology, 24*, 776–781.
Camras, L. A., Sachs-Alter, E., & Ribordy, S. C. (1996). Emotion understanding in maltreated children: Recognition of facial expressions and integration with other emotion cues. In M. Lewis & M. Sullivan (Eds.), *Emotional development in atypical children* (pp. 203–225). Hillsdale, NJ: Erlbaum.
Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child Development*, *59*, 604–616.
Carver, L. J., Dawson, G., Panagiotides, H., Meltzoff, A. N., McPartland, J., Gray, J., & Munson, J. (2003). Age-related differences in the neural correlates of face recognition during the toddler and preschool years. *Developmental Psychobiology, 42*, 148–159.
Darwin, C. (2002). *The expression of emotions in man and animals*. New York: Oxford University Press (Original work published 1872).
DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 1174–1176.
deGelder, B., Böcker, K. B. E., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion in voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, *260*, 133–136.
deGelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*, 289–311.
Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.
Ekman, P. (1994). Strong evidence for universals in facial expression: A reply to Russell's mistaken critique. *Psychological Bulletin*, *115*, 268–287.
Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, *64*, 657–674.
Fingelkurts, A. A., Fingelkurts, A., Krause, C. M., Moettoenen, R., & Sams, M. (2003). Cortical operational synchrony during audio-visual speech integration. *Brain and Language*, *85*, 297–312.
Fox, N. A. (1994). The development of emotion regulation: Biological and behavioral considerations. *Monographs of the Society for Research in Child Development*, *59* (2–3, Serial No. 240).
Hainline, L., & Abramov, I. (1992). Assessing visual development: Is infant vision good enough. *Advances in Infancy Research*, *7*, 39–102.
Hayes, D. S., & Kelly, S. B. (1984). Young children's processing of television: Modality differences in the retention of temporal relations. *Journal of Experimental Child Psychology, 38*, 505–514.
Hayes, D. S., Kelly, S. B., & Mandel, M. (1986). Media difference in children's story synopses: Radio and television contrasted. *Journal of Educational Psychology, 78*, 341–346.
Hess, U., Kappas, A., & Scherer, K. R. (1988). Multichannel communication of emotion: Synthetic signal production. In K. R. Scherer (Ed.), *Facets of emotion: Recent research* (pp. 161–182). Hillsdale, NJ: Erlbaum.
Hollingshed, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.
Hölscher, C., Rolls, E. T., & Xiang, J.-Z. (2003). Perirhinal cortex neuronal activity related to long-term familiarity memory in the macaque. *European Journal of Neuroscience*, *18*, 2037–2046.
Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry, 43*, 280–283.
Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Development*, *72*, 352–369.
Klinnert, M. D., Emde, R. N., Butterfield, P., & Campos, J. J. (1986). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. *Developmental Psychology, 22*, 427–432.
Lewkowicz, D. J. (1988). Sensory dominance in infants: 1. Six-month-old infants' response to auditory–visual compounds. *Developmental Psychology, 24*, 155–171.
Lewkowicz, D. J. (1996). Perception of auditory–visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1094–1106.
Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review, 3*, 215–221.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.
Montague, D. P. F., & Walker-Andrews, A. S. (2002). Mothers, fathers, and infants: The role of person familiarity and parental involvement in infants' perception of emotion expressions. *Child Development*, *73*, 1339–1352.
Napolitano, A. V., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, *75*, 1850–1870.
Pezdek, K., & Stevens, E. (1984). Children's memory for auditory and visual information on television. *Developmental Psychology, 20*, 212–218.
Pollak, S. D., Klorman, R., Thatcher, J. E., & Cicchetti, D. (2001). P3b reflects maltreated children's reactions

to facial displays of emotion. *Psychophysiology, 28*, 267–274.

Pollak, S. D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology, 38*, 784–791.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*, 1387–1401.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*, 76–92.

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, *15*, 123–148.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for the auditory modality in young children. *Child Development*, *74*, 822–833.

Soken, N. H., & Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Development*, *63*, 787–795.

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year olds. *Developmental Psychology, 21*, 195–200.

Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the Parent–Child Conflict Tactics Scales: Development and psychometric data for a national sample of American parents. *Child Abuse and Neglect*, *22*, 249–270.

Trickett, P. K., Aber, J. L., Carlson, V., & Cicchetti, D. (1991). Relationship of socioeconomic status to the etiology and developmental sequelae of physical child abuse. *Developmental Psychology, 27*, 148–158.

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