Developmental Science 7:3 (2004), pp 261–269

REPORT

Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior

Amrisha Vaish and Tricia Striano

Department of Cultural Ontogeny, Max Planck Institute for Evolutionary Anthropology, Germany

Abstract

To examine the influences of facial versus vocal cues on infants' behavior in a potentially threatening situation, 12-month-olds on a visual cliff received positive facial-only, vocal-only, or both facial and vocal cues from mothers. Infants' crossing times and looks to mother were assessed. Infants crossed the cliff faster with multimodal and vocal than with facial cues, and looked more to mother in the Face Plus Voice compared to the Voice Only condition. The findings suggest that vocal cues, even without a visual reference, are more potent than facial cues in guiding infants' behavior. The discussion focuses on the meaning of infants' looks and the role of voice in development of social cognition.

Introduction

Social referencing is a process of communication whereby people actively seek and use others' perceptions and interpretations of ambiguous situations to form their own interpretations of those situations (Feinman, 1982; Sorce, Emde, Campos & Klinnert, 1985). Typical social referencing paradigms involve exposing infants to novel, ambiguous situations, such as the 'visual cliff' (a Plexiglas surface providing invisible support over an apparent drop; e.g. Sorce *et al.*, 1985), strangers (e.g. Clarke-Stewart, 1978; Feinman & Lewis, 1983), or animated and noisy toys (the 'novel toy paradigm'; e.g. Hornik, Risenhoover & Gunnar, 1987). Social referencing studies demonstrate that infants guide their behavior based on the emotional cues people offer them (e.g. Camras & Sachs, 1991; Klinnert, Emde, Butterfield & Campos, 1986).

Many of these studies focus on the impact of positive or negative valence (e.g. Boccia & Campos, 1983; Bradshaw, Campos & Klinnert, 1986; Klinnert, 1984; Sorce *et al.*, 1985). Adults are often asked to convey emotional information to infants through multiple communication channels (e.g. Gunnar & Stone, 1984; Hirshberg & Svejda, 1990; Hornik *et al.*, 1987). However, multimodal cues create a confounding effect, blurring the weight of each channel in the social referencing process (Gewirtz & Peláez-Nogueras, 1992). Given the endless array of information available to infants, it is important to consider which channels of communication infants use.

Controlled signals conveyed through only one channel of communication have been used in some studies. However, most of this research has been limited to the use of facial expressions (e.g. Feinman, 1992; Klinnert, 1984; Sorce *et al.*, 1985; Zarbatany & Lamb, 1985), leading to social referencing being considered a primarily visual phenomenon (Recchia, 1997). In a classic visual cliff study (Sorce *et al.*, 1985), mothers posed facial-only expressions of fear, anger, interest and joy. Most infants crossed the cliff if mothers expressed interest or joy, but few crossed if they expressed fear or anger.

Communication channels other than the face have received limited attention (Cohen, DeLoache & Strauss, 1979; Eimas, 1975). This is surprising considering that infants respond appropriately to other modes of communication, especially the voice, from early months (see Bühler & Hetzer, 1928; Charlesworth & Kreutzer, 1973; Fernald, 1992; Soken & Pick, 1999; Walker-Andrews & Gronlick, 1983). In fact, vocally expressed emotions may be stronger signals than facial ones because they carry acoustic properties that can directly engender emotions in infants (Fernald, 1993). Furthermore, infants often experience vocal-only cues, such as when a parent talks from behind the infant. The vocal channel is therefore apt to be especially effective in emotional communication (Baldwin & Moses, 1996;

Address for correspondence: Tricia Striano, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany; e-mail: striano@eva.mpg.de

Clyman, Emde, Kempe & Harmon, 1986; Emde, 1992; Feinman, 1992; Sorce *et al.*, 1985), but its effectiveness in social referencing needs empirical assessment.

Only one published study involved such an assessment. Mumme, Fernald and Herrera (1996) examined the effects of facial and vocal cues on 12-month-olds' behavior toward novel toys. In the face-only condition, the mother looked at the infant and toy while displaying positive, neutral or negative facial expressions; in the voice-only condition, she had her back to the infant but could see the toy, and vocally expressed one of the three expressions. Infants responded appropriately to fearful vocal cues – they looked longer to mother, and displayed less toy proximity and more negative affect – but not to facial cues, suggesting that the voice alone, but not the face alone, is powerful enough to guide infants' behavior.

These results contrast with Sorce *et al.*'s (1985) finding that the face is a potent behavior modifier. One possible explanation for this discrepancy is that the visual cliff is a dangerous situation in which infants use facial-only cues to modify their actions, whereas in ambiguous but non-threatening situations such as the novel toy paradigm, infants additionally require vocal cues (see also Hirshberg & Svejda, 1990; Mumme *et al.*, 1996). Since Sorce *et al.* (1985) did not assess the effect of vocal cues alone, however, this hypothesis needs empirical examination.

The present study explores the effects of facial versus vocal cues on 12-month-olds in a potentially threatening situation. Infants on the shallow end of the visual cliff received facial and vocal cues ('Face Plus Voice'), facial cues only ('Face Only'), or vocal cues only ('Voice Only') from their mothers. This is the first visual cliff study with a condition (Voice Only) in which mothers were not looking at the cliff.

Based upon findings that infants best recognize messages conveyed through the face and voice (e.g. Bahrick, 1992, 1994; Lewkowicz, 1996; Walker-Andrews, 1997), we predicted that infants would cross the cliff faster in Face Plus Voice than in the other conditions. Additionally, we predicted that if visual reference is necessary for social referencing (Moses, Baldwin, Rosicky & Tidball, 2001; Striano & Rochat, 2000), infants would cross faster in Face Only than in Voice Only. If, however, the vocal channel more readily influences infants' behavior (e.g. Mumme *et al.*, 1996), infants would cross faster in Voice Only compared to Face Only.

Methods

Participants

Eighty-nine full-term infants were recruited from a list of parents who had volunteered to participate in child development studies. Of these 89, 49% ($n = 44$) were excluded (comparable to the 40% excluded in Sorce *et al.*, 1985) for the following reasons: 22.7% because mothers did not follow directions ($n = 10$; 7 because mothers gestured, 2 because they talked prior to being signalled and 1 because she insisted on placing the infant on the cliff), 22.7% due to experimenter error ($n = 10$; 5 due to imprecise set-up, 3 because sessions were inappropriately recorded, 1 because the mother was incorrectly instructed and 1 because the infant was initially placed with his feet over the cliff), 38.6% because infants became fussy ($n = 17$; 7 were fussy prior to being on the cliff, and 10 became fussy after), 11.4% due to equipment failure ($n = 5$), 2.3% because the infant crossed without noticing the cliff ($n = 1$) and 2.3% because the infant crossed without referencing ($n = 1$). Of the 45 infants in the final sample (24 females; *M*: 12 months, 2 days, SD: 11 days; range: 11 months, 14 days to 12 months, 27 days), 14 were in each of the three conditions, and 3 were in a fourth condition which was discontinued (see procedure).

Setting

The experiment took place in a laboratory room equipped with a visual cliff (a 195×94.5 cm Plexiglas-covered table divided into a shallow half under which a chequered surface is placed immediately beneath the Plexiglas, and a deep half under which a similar chequered surface is placed some variable distance beneath the Plexiglas). The height of the cliff was 28 cm.[1] Three digital video cameras filmed the infant, the mother's face and torso, and a panoramic view of the cliff (see Figure 1). All images were synchronized with a quad splitter and recorded on a mini-VCR, and were displayed live on a TV screen that mothers watched during the experiment.

A curtain initially hid the visual cliff when the participants entered the laboratory room. This ensured that when the infant was placed on the visual cliff, she would experience an entirely novel situation, and also that the mother did not cue the infant about the cliff before the experiment began. Once the infant was on the cliff, the curtain was drawn by Experimenter 1 (E1), but only enough so that the mother, and not the TV screen, was visible to the infant.

Procedure

A third of the participants were assigned to Face Plus Voice, Face Only, and Voice Only each. We tested three

[1] This was the closest depth possible on our visual cliff to the 30-cm depth used by Sorce *et al.* (1985) to create an ambiguous situation for 12-month-olds.

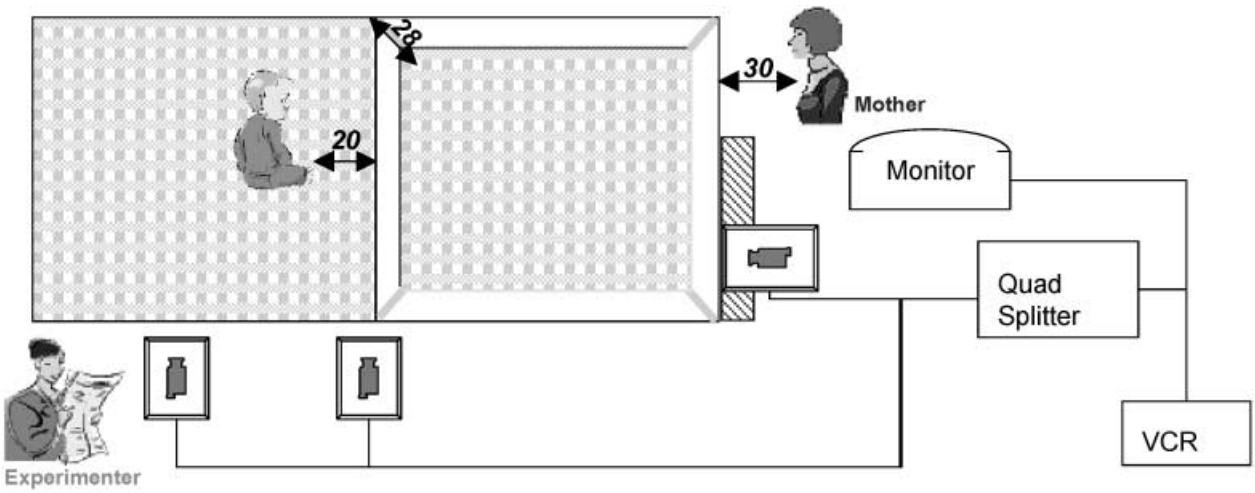

Figure 1 *Experimental set-up.*

infants in a condition in which mothers looked away and were silent ('No Cues'). Because all three infants became fussy, and because our aim was specifically to compare facial versus vocal cues, this condition was not deemed necessary, and was aborted.

Prior to testing, mothers were told that they could stop the experiment if they sensed that their infants were agitated. If infants became fussy as a second experimenter (E2) took them from their mothers, they were allowed to stay with the mothers until they were calmer. E2 then re-attempted to put them on the cliff, but aborted the experiment if they became fussy again. Also, if infants became fussy once on the cliff, the experimenters waited for one minute, but aborted the experiment if the infants did not calm down.

Mothers stood approximately 30 cm beyond the deep side of the cliff with their back to the cliff and watched on a TV screen as E2 placed the infants on the shallow side with their feet approximately 20 cm from the 'drop',[2] at which time E1 drew the curtain. E2 stood next to the cliff throughout the experiment to ensure infants' safety, but pretended to read a magazine so as not to provide cues. E1 now observed infants on the mini-VCR screen, and when infants looked up to mother after looking down at the 'drop', she showed mothers a sign that read, 'You can start!'[3] Mothers began positively cueing infants to cross according to one of the following conditions:

Face Plus Voice. Faced the cliff, smiled, and vocalized.
Face Only. Faced the cliff, and only smiled and nodded (i.e. did not vocalize).
Voice Only. Did not face the cliff, continued watching TV screen, and vocalized.

Mothers were requested not to gesture at all during the experiment.

Coding

Two behaviors were coded as follows:

Crossing Time

Duration in seconds for infants to cross the cliff. Timing started when mothers started cueing infants and ended the moment the infants' entire body had crossed the cliff.[4]

Infant Looking

Duration of looking to mothers' head or face as a proportion of Crossing Time and number of such looks during Crossing Time.

Reliability

A blind coder first watched 20% of the videos to ensure that E1 had signalled mothers to cue only once infants had looked down at the cliff and then looked to mother. Agreement was 100%. A blind coder also coded Crossing

[2] This distance was chosen because it was close enough to the 'drop' to make the infant notice it soon after being placed on the cliff, but far enough away to create an uncertain situation that she would need to actively explore and reference about.

[3] This criterion was based upon past work on social referencing (e.g. Klinnert, 1984; Sorce *et al.*, 1985), in which mothers could begin cueing only once the infant had produced a true 'referencing' look, i.e. looked to the adult only after looking at the ambiguous situation.

[4] Unlike Sorce *et al.*'s (1985) criterion that infants had to reach the end of the cliff, our end-point was defined based upon the assumption that infants had overcome their wariness of the apparent cliff when they had crossed it with their entire body.

Time and Infant Looking for these sessions. Pearson Correlation for Crossing Time between coders was 1.0, $p < .0001$, and Cohen's kappa on Infant Looking time was .84. Since mothers were not instructed on what to say to infants, and may have vocalized more when not providing facial cues (i.e. in Voice Only), amount of mothers' vocalizations in 20% of Face Plus Voice and Voice Only sessions was also coded. A one-way ANOVA (two levels: Face Plus Voice; Voice Only) revealed no significant difference, $p = .506$.

Results

Crossing Time

Crossing Times were analyzed using a 2 (sex) × 3 (condition) between-subjects ANOVA, which revealed no significant effects of sex, $p = .337$, and no significant sex × condition effects, $p = .707$. There was a significant condition effect, $F(2, 39) = 5.96$, $p = .009$. Pair-wise comparisons indicated lower Crossing Time in Face Plus Voice ($M = 58.63$, SD = 47.11) than in Face Only ($M = 215.74$, SD = 179.30, $p = .003$), as well as in Voice Only ($M = 102.97$, SD = 108.87) than in Face Only ($M = 215.74$, SD = 179.30, $p = .036$). However, Crossing Times across Face Plus Voice and Voice Only were *not* different, $p = .28$ (see Figure 2).

To ensure that the high standard deviations of the data were not due to a few extreme values, *z*-scores were calculated for all infants. All *z*-values were below 2.65 (the standard *z*-value used to determine whether a score in a sample of 14 is an outlier), indicating that none of the scores were unusually far from the mean (Tabachnick & Fidell, 2001). To confirm the ANOVA results, a Kruskal-Wallis non-parametric test was also conducted on Crossing Times. This test also revealed a significant overall condition difference, $\chi^2(2, N = 42) = 7.83$, $p = .02$. Pair-wise comparisons, conducted using Mann-Whitney *U* tests, supported the results from the ANOVA pair-wise comparisons. Crossing Times were different across Face Plus Voice ($M = 58.63$, SD = 47.11) and Face Only ($M = 215.74$, SD = 179.30, $z = -2.80$, $p = .004$), and approached significance across Voice Only ($M = 102.97$, SD = 108.87) and Face Only ($M = 215.74$, SD = 179.30, $z = -1.84$, $p = .069$). Confirming the ANOVA results, Crossing Times were not significantly different across Face Plus Voice and Voice Only, $p = .541$.

Infant Looking

Duration of looking

A one-way ANOVA (three levels: Face Plus Voice; Face Only; Voice Only) was conducted, which revealed a condition effect that approached conventional significance, $F(2, 39) = 3.07$, $p = .058$ (see Figure 3). Pair-wise comparisons revealed that infants looked to mother more in Face Plus Voice ($M = 30.47$, SD = 15.32) than in Voice Only ($M = 17.29$, SD = 9.34, $p = .019$). Amount of looking was not different across Face Plus Voice and Face Only ($M = 22.34$, SD = 16.85, $p = .138$), nor across Voice Only and Face Only, $p = .352$.

Number of looks

A one-way ANOVA (three levels: Face Plus Voice; Face Only; Voice Only) revealed no significant differences for proportional numbers of infants' looks across conditions, $p = .472$.

Discussion

To investigate what cues infants use when they encounter a potentially threatening situation, we placed 12-month-

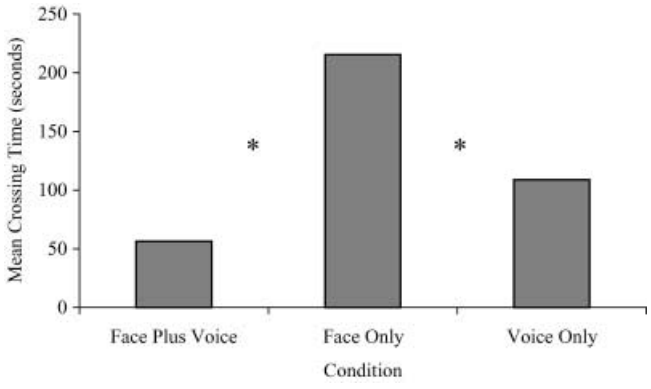

Figure 2 *Mean times for infants to cross the visual cliff as a function of condition.*

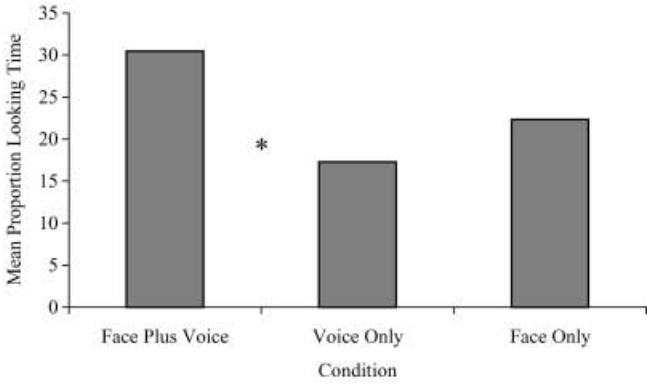

Figure 3 *Mean time infants in the three conditions spent looking to mother as a proportion of Crossing Times.*

old infants on a visual cliff and manipulated the information that mothers provided. Infants were presented positive facial and vocal cues, facial-only cues, or vocal-only cues. Infants crossed faster in response to vocal-only than to facial-only cues. This finding is intriguing considering that mothers did not have visual access to the cliff in the Voice Only condition, and suggests that the voice is a more potent channel of emotional communication than the face (see also Baldwin & Moses, 1996; Feinman, 1992; Fernald, 1992).

Interestingly, infants did cross the cliff even in Face Only. We therefore agree with Hirshberg and Svejda (1990) and Mumme *et al.* (1996) that infants may not use facial-only cues to guide their behavior in ambiguous situations, but may do so in potentially threatening situations. Accordingly, infants tested in the novel toy paradigm do not consistently rely on others' facial signals (e.g. Klinnert, 1984; Mumme *et al.*, 1996; Zarbatany & Lamb, 1985), whereas infants tested on the visual cliff do (Sorce *et al.*, 1985). The current study confirms this finding. However, while facial-only cues were sufficient to make infants cross the cliff, vocal-only cues were more effective in doing so, raising the issue of what makes the voice a more powerful modifier of infant behavior.

One possibility is that infants commonly receive vocal cues without accompanying facial cues. Infants may therefore learn to respond appropriately to vocal-only cues, and to trust the voice so that they do not also need facial cues in order to appropriately social reference (Baldwin & Moses, 1996). Accordingly, mothers in Voice Only did not vocalize more; thus, the only difference between the information infants received in Face Plus Voice versus Voice Only was the facial cues. That infants crossed equally fast in these two conditions suggests that the facial cues in Face Plus Voice did not significantly increase the potency of the signal (see also Lewkowicz, 1988; Walker-Andrews, 1997).

Another possibility is that unlike facial cues, infants need not actively seek out the voice in order to gather information from it (Feinman, 1985). Furthermore, when receiving vocal-only information, infants can visually focus on and assess the novel situation while simultaneously gathering auditory information about how to react. With facial-only cues, however, infants must alternate their visual attention between the face and the object, thus considerably slowing the referencing process. A revealing behavior to examine in this regard was that of infants' looks to mother, since these provided facial but not vocal information.

One counterintuitive finding was that infants looked to mother equally in Face Only and Voice Only. Since infants in Voice Only were not receiving visual information, it seems surprising that they looked as much as those infants receiving *only* visual information. However, past research shows that infants look to a speaker even when the speaker is communicating only vocally (e.g. Hornik & Gunnar, 1988; Svejda & Campos, 1982; Walden & Ogan, 1988), and that they orient to the face of the speaker (Carr, Dabbs & Carr, 1975; Clyman *et al.*, 1986).

Feinman, Roberts, Hsieh, Sawyer and Swanson (1992) suggest that these looks to speaker mean that infants are seeking additional information. However, since infants in our study responded faster in Voice Only than in Face Only, they were likely not seeking additional information through their looks in Voice Only. This is also supported by the finding that infants looked more in Face Plus Voice than in the other conditions. Infants in Face Plus Voice were receiving information through *two* channels of communication; if infants' looks were indeed intended to gather additional information, infants should have looked up *less* in the multimodal compared to the unimodal conditions.

To understand the meanings of infants' looks, it is important to assess *types* of infants' looks (e.g. Hornik & Larson, 1988), possibly by using Clyman *et al.*'s (1986) typology, in which eight social looks are identified. Clyman *et al.* found that 'social referencing looks' (defined as occurring after the infant attends to an ambiguous event but before she acts upon it) occurred least frequently of all looks. In our study, all infants produced these looks at least once, since it was only after such looks that mothers began cueing. However, all subsequent looks to mother were not necessarily social referencing looks. While a detailed analysis of infants' looks is beyond the scope of this paper, further research is clearly needed to assess infants' looks in referencing situations.

Interestingly, the condition differences in proportion of looking time were not replicated in number of looks. This discrepancy fits Walden and Baxter's (1989) hypothesis that different indices of social looking may have important functional differences, and should not be generalized into one 'looking' measure. While it is unlikely that one index measures only one underlying component, a broad distinction made between indices and tested empirically can provide exciting answers about the nature of infants' looks.

That infants crossed in Voice Only contrasts with past work, which generally presumes that a person's emotions concern the object of her gaze (Woodward, 2003; Moses *et al.*, 2001; Poulin-Dubois, 1999; Phillips, Wellman & Spelke, 2002). Indeed, Striano and Rochat (2000) found that 10-month-olds reference about a novel toy more when the experimenter is looking at them than away. Similarly, 12-month-olds use an adult's gaze direction to determine which object to link the adult's emotional outbursts to (Moses *et al.*, 2001). Such research suggests

that infants need gaze direction to deduce the object of a referee's communication (see also Baldwin, 1991, 1993; Dunham, Dunham & Curwin, 1993; Mumme & Fernald, 2003; Repacholi, 1998), whereas our findings suggest that gaze direction is not so essential. This discrepancy might be due to the fact that both Striano and Rochat (2000) and Moses *et al.* (2001) used novel toys, which present ambiguous but harmless situations, whereas we used the visual cliff, which is a threatening situation. It is possible that when infants feel threatened, they use adults' cues even if those cues do not appear to be *visually* referential.

This does not, however, imply that mood modification regulates infants' behavior in threatening situations (e.g. Feinman, 1982). Rather, when infants cannot determine the referee's gaze-direction, they nevertheless resourcefully use vocal-only cues to modify their behavior, supporting Campos and Stenberg's (1981) hypothesis that infants will increasingly use vocal cues when the mother is visually inaccessible. In this context, it is necessary to consider 'instrumental' versus 'affective' social referencing. 'Affective' messages show infants what to feel, whereas 'instrumental' messages explicitly convey what infants should do (Campos, 1983; Feinman, 1982, 1983; Hornik & Gunnar, 1988; Klinnert, Campos, Sorce, Emde & Svejda, 1983). In our study, in Face Only, mothers could only provide affective cues (e.g. smiling), whereas in Voice Only, they could convey affective messages through their tone of voice and motherese, *and* could give instrumental instructions, such as to take a step forward, cross the cliff, etc.

Instrumental referencing might be more potent than affective referencing (Feinman *et al.*, 1992; Hornik & Gunnar, 1988; Klinnert *et al.*, 1983). Only when infants learn how to act upon a stimulus can they feel a sense of control, and thus be less fearful or ambivalent about the stimulus (Bandura, 1977; Hornik-Parritz, Mangelsdorf & Gunnar, 1992). Accordingly, Hornik and Gunnar (1988) found that although mothers' affective messages helped infants overcome their fear of a caged rabbit enough to approach it, only instrumental messages (i.e. mothers touching the rabbit) made infants brave enough to touch the rabbit. Consistently, our finding that infants crossed faster in the two Voice conditions than in Face Only suggests that instrumental social referencing might indeed be more powerful, at least in threatening situations. At a minimum, it suggests that a combination of affective and instrumental messages is more powerful than affective messages alone. Perhaps the voice is such a powerful mode of communication precisely because it is capable of providing both instrumental and affective messages.

Interestingly, Mumme *et al.* (1996) controlled for this difference between face and voice. Mothers in the voice-only condition said phrases meaningless to infants but with distinct prosody (thus conveying only affect). Infants nevertheless responded to the voice but not the face, suggesting that in the current study, it was not the voice's instrumental component alone that made infants cross faster.

While face-to-face interactions have been emphasized in the development of social cognition (Striano & Rochat, 1999), little emphasis has been placed on other modes of communication (Recchia, 1997). However, reliable mechanisms of development should be experienced by infants across cultures (Tomasello, 1995), and have deep evolutionary roots. Vocal, but not facial, information meets these criteria. A sensitivity to vocal cues is well established from birth (Mastropieri & Turkewitz, 1999), and infants can experience vocal but not facial cues while being carried on the back or from a distance.

Research points to cross-cultural differences in how much face-to-face contact infants experience. In hunter-gatherer cultures (Lozoff, Brittenham, Trause, Kennell & Klaus, 1977) and cultures in parts of Guatemala (Mata, 1978), young infants are usually carried on caregivers' backs. In the case of blind infants, there is no face-to-face contact whatsoever (Hobson, 1993; Bigelow, 1995). In such cultures, caregivers' vocalizations are likely crucial for infants to share attention and communicate (Koester & Traci, 1999; Recchia, 1997).

Studies with animals also reveal the importance of the auditory channel. Blind infant crab-eating monkeys use auditory signals to maintain social affinity and closeness to mothers (Berkson & Becker, 1975). Barbary macaques substitute vocalizations for visual signals when visual signals are impaired, such as during close body contact (Kipper & Todt, 2002). Finally, vervet monkeys use far more vocal than visual signals, possibly because the former can be used even without face-to-face contact (Cheney & Seyfarth, 1985, 1990). Clearly, sensitivity to vocal cues is instrumental, works in a wider range of contexts than facial cues, and is therefore a good mechanism of information communication.

The present study confirms that the human voice is a deep-rooted and highly powerful modifier of infant behavior not only in ambiguous situations (Mumme *et al.*, 1996), but also in potentially threatening situations. Sensitivity to vocal cues is a highly adaptive skill, deserving much more attention in the ontogeny of social referencing and human social cognition in general.

Acknowledgements

We are grateful to the Cultural Ontogeny group, and especially to Carola Arfeller, Didrik Drageset, Anne Henning,

Jennifer Kittel, Saskia Stieber and Kerstin Träger for help with participant recruitment and with the running of the study, and to Alexander Burkhardt for technical assistance and graphics. Thank you to Michael Tomasello for comments on the paper. We also thank two anonymous reviewers for helpful feedback, the Universitätsfrauenklinik and the Eitingon Krankenhaus for help with participant recruitment, and the parents and infants who participated.

References

Bahrick, L.E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology*, **53**, 180–199.

Bahrick, L.E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology*, **6**, 111–123.

Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, **63**, 875–890.

Baldwin, D.A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, **20**, 395–418.

Baldwin, D.A., & Moses, L.J. (1996). The ontogeny of social information gathering. *Child Development*, **67**, 1915–1939.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, **84**, 191–215.

Berkson, G., & Becker, J.D. (1975). Facial expressions and social responsiveness of blind monkeys. *Journal of Abnormal Psychology*, **84** (5), 519–523.

Bigelow, A.E. (1995). The effect of blindness on the early development of the self. In P. Rochat (Ed.), *The self in infancy: Theory and research, 112* (pp. 327–347). Amsterdam: Elsevier North Holland.

Boccia, M.L., & Campos, J.J. (1983). Maternal emotional signals and infants' reactions to strangers. Paper presented at the biennial meeting of the Society for Research in Child Development, Detroit (April).

Bradshaw, D.L., Campos, J.J., & Klinnert, M.D. (1986). Emotional expressions as determinants of infants' immediate and delayed responses to prohibitions. Paper presented at the Fifth International Conference on Infant Studies, Los Angeles (April).

Bühler, C., & Hetzer, H. (1928). Das erste Verständnis für Ausdruck im ersten Lebensjahr. *Zeitschrift für Psychologie*, **107**, 50–61.

Campos, J.J. (1983). The importance of affective communication in social referencing: a commentary on Feinman. *Merrill-Palmer Quarterly*, **29**, 83–87.

Campos, J.J., & Stenberg, C. (1981). Perception, appraisal, and emotion: the onset of social referencing. In M.E. Lamb & L.R. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 273–314). Hillsdale, NJ: Erlbaum.

Camras, L.A., & Sachs, V.B. (1991). Social referencing and caretaker expressive behavior in a day care setting. *Infant Behavior and Development*, **14**, 27–36.

Carr, S.J., Dabbs, J.M., & Carr, T.S. (1975). Mother–infant attachment: the importance of the mother's visual field. *Child Development*, **46**, 331–338.

Charlesworth, W.R., & Kreutzer, M.A. (1973). Facial expressions of infants and children. In P. Ekman (Ed.), *Darwin and facial expressions: A century of research in review* (pp. 91–168). New York: Academic Press.

Cheney, D.L., & Seyfarth, R.M. (1985). Vervet monkey alarm calls: manipulation through shared information? *Behaviour*, **94**, 150–166.

Cheney, D.L., & Seyfarth, R.M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Clarke-Stewart, K.A. (1978). Recasting the lone stranger. In J. Glick & K.A. Clarke-Stewart (Eds.), *The development of social understanding* (pp. 109–176). New York: Gardner Press.

Clyman, R.B., Emde, R.N., Kempe, J.E., & Harmon, R.J. (1986). Social referencing and social looking among twelve-month-old infants. In T.B. Brazelton & M.W. Yogman (Eds.), *Affective development in infancy* (pp. 75–94). Norwood, NJ: Ablex.

Cohen, L.B., DeLoache, J.S., & Strauss, M.S. (1979). Infant visual perception. In J.D. Osofsky (Ed.), *Handbook on infant development* (pp. 393–438). New York: Wiley.

Dunham, P.J., Dunham, F., & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology*, **29**, 827–831.

Eimas, P.D. (1975). Speech perception in early infancy. In L.B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition. Perception of space, speech, and sound* (Vol. 2, pp. 193–231). New York: Academic Press.

Emde, R.N. (1992). Social referencing research: uncertainty, self, and the search for meaning. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 79–94). New York: Plenum.

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly*, **28** (4), 445–470.

Feinman, S. (1983). How does baby socially refer? Two views of social referencing: a reply to Campos. *Merrill-Palmer Quarterly*, **29**, 467–471.

Feinman, S. (1985). Emotional expression, social referencing, and preparedness for learning in infancy – mother knows best, but sometimes I know better. In G. Zivin (Ed.), *The development of expressive behavior: Biology–environment interactions* (pp. 291–318). New York: Academic Press.

Feinman, S. (1992). In the broad valley: an integrative look at social referencing. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 3–13). New York: Plenum.

Feinman, S., & Lewis, M. (1983). Social referencing at 10 months: a second-order effect on infants' responses to strangers. *Child Development*, **54**, 878–887.

Feinman, S., Roberts, D., Hsieh, K., Sawyer, D., & Swanson, D. (1992). A critical review of social referencing in infancy. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 15–54). New York: Plenum.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford: Oxford University Press.

Fernald, A. (1993). Approval and disapproval: infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, **64**, 657–674.

Gewirtz, J.J., & Peláez-Nogueras, M. (1992). Social referencing as a learned process. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 151–173). New York: Plenum.

Gunnar, M.R., & Stone, C. (1984). The effects of positive maternal affect on infant responses to pleasant, ambiguous, and fear-provoking toys. *Child Development*, **55**, 1231–1236.

Hirshberg, L.M., & Svejda, M. (1990). When infants look to their parents: I. Infants' social referencing of mothers compared to fathers. *Child Development*, **61**, 1175–1186.

Hobson, R.P. (1993). *Autism and the development of the mind*. Hove, East Sussex: Lawrence Erlbaum.

Hornik, R., & Gunnar, M.R. (1988). A descriptive analysis of infant social referencing. *Child Development*, **59**, 626–634.

Hornik, R., & Larson, M. (1988). Social referencing: relations between temperament and patterns of looking. Paper presented at the International Conference on Infant Studies, Washington, DC.

Hornik, R., Risenhoover, N., & Gunnar, M.R. (1987). The effects of maternal positive, neutral, and negative affective communications on infant responses to new toys. *Child Development*, **58**, 937–944.

Hornik-Parritz, R., Mangelsdorf, S., & Gunnar, M.R. (1992). Control, social referencing, and the infant's appraisal of threat. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 209–228). New York: Plenum.

Kipper, S., & Todt, D. (2002). The use of vocal signals in the social play of Barbary macaques. *Primates*, **43** (1), 3–17.

Klinnert, M.D. (1984). The regulation of infant behavior by maternal facial expression. *Infant Behavior and Development*, **7**, 447–465.

Klinnert, M.D., Campos, J.J., Sorce, J.F., Emde, R.N., & Svejda, M. (1983). Emotions as behavior regulators: social referencing in infancy. In R. Plutchik & H. Kelleman (Eds.), *Emotion: Theory, research and experience: Vol. 2. Emotions in early development* (pp. 57–86). New York: Academic Press.

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.

Koester, L.S., & Traci, M.A. (1999). Cultural variations in mother–infant gaze behavior. Paper presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM (April).

Lewkowicz, D.J. (1988). Sensory dominance in infants: 1. Six-month-old infants' response to auditory-visual compounds. *Developmental Psychology*, **24** (2), 155–171.

Lewkowicz, D.J. (1996). Infants' response to the audible and visible properties of the human face, I: Role of lexical/syntactic content, temporal synchrony, gender, and manner of speech. *Developmental Psychology*, **32**, 347–366.

Lozoff, B., Brittenham, G., Trause, M.K., Kennell, J., & Klaus, M. (1977). The mother–newborn relationship: limits of adaptability. *Journal of Pediatrics*, **91** (1), 1–12.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*, **35**, 204–214.

Mata, L.J. (1978). *The children of Santa Maria Cauque: A prospective field study of health and growth*. Cambridge, MA: MIT Press.

Moses, L.J., Baldwin, D.A., Rosicky, J.G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development*, **72**, 718–735.

Mumme, D.L., & Fernald, A. (2003). The infant as onlooker: learning from emotional reactions observed in a television scenario. *Child Development*, **74**, 221–237.

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.

Phillips, A.T., Wellman, H.M., & Spelke, E.S. (2002). Infants' ability to connect gaze and emotional expression to intentional actions. *Cognition*, **85**, 53–78.

Poulin-Dubois, D. (1999). Infants' distinction between animate and inanimate objects: the origins of naive psychology. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 257–280). Mahwah, NJ: Erlbaum.

Recchia, S.L. (1997). Social communication and response to ambiguous stimuli in toddlers with visual impairments. *Journal of Applied Developmental Psychology*, **18**, 297–316.

Repacholi, B.M. (1998). Infants' use of attentional cues to identify the referent of another person's emotional expression. *Developmental Psychology*, **34**, 1017–1025.

Soken, N.H., & Pick, A.D. (1999). Infants' perception of dynamic affective expressions: do infants distinguish specific expressions? *Child Development*, **70**, 1275–1282.

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

Striano, T., & Rochat, P. (1999). Developmental link between dyadic and triadic social competence in infancy. *British Journal of Developmental Psychology*, **17**, 551–562.

Striano, T., & Rochat, P. (2000). Emergence of selective social referencing in infancy. *Infancy*, **1** (2), 253–264.

Svejda, M.J., & Campos, J.J. (1982). Mother's vocal expression of emotion as a behavior regulator. Paper presented at the Third International Conference on Infant Studies, Austin, TX (March).

Tabachnick, B.G., & Fidell, L.S. (2001). *Computer-assisted research design and analysis*. Boston: Allyn and Bacon.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P.J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum.

Walden, T.A., & Baxter, A. (1989). The effect of context and age on social referencing. *Child Development*, **60**, 1511–1518.

Walden, T.A., & Ogan, T.A. (1988). The development of social referencing. *Child Development*, **59**, 1230–1240.

Walker-Andrews, A.S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological Bulletin*, **121** (3), 1–20.

Walker-Andrews, A.S., & Gronlick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development*, **6**, 491–498.

Woodward, A.L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, **6** (3), 297–311.

Zarbatany, L., & Lamb, M.E. (1985). Social referencing as a function of information source: mothers versus strangers. *Infant Behavior and Development*, **8**, 25–33.

Received: 13 May 2003
Accepted: 3 July 2003