Contribution of facial and vocal cues in the still-face response of 4-month-old infants

Tricia Striano*, Evelin Bertin

Max Planck Institute for Evolutionary Anthropology, Cultural Ontogeny, Deutscher Platz 6, 04103 Leipzig, Germany

Received 26 February 2004; received in revised form 14 June 2004; accepted 19 June 2004

Abstract

The contribution of contingent facial and vocal information in the still-face effect was investigated. Four-month-old infants either saw and heard their mother, only saw their mother, or only heard their mother interacting with them. These interaction periods were followed by the cessation of the mother’s interactive face and/or voice. Only infants who observed their mother’s face become still and neutral, showed a still-face effect by decreasing their visual attention and positive affect. The findings provide further support that the mother’s interactive voice does not contribute to the still-face effect. The developing sensitivity to vocal information in dyadic and triadic contexts is discussed.

Keywords: Still-face; Vocal cues; Facial cues; Infant social development

1. Introduction

From as early as 2 months of age, human infants reliably express social expectations in relation to people. One of the most reliable paradigms to assess what infants understand and expect from other people is the still-face paradigm (Adamson & Frick, 2003). In this paradigm, an adult, often the infant’s mother or sometimes a stranger, interacts with the infant in a normal face-to-face interaction. Then, the adult suddenly freezes, becomes unresponsive, and poses a stationary neutral face. This still-face period is typically followed by another normal face-to-face interaction. The whole interaction episode usually lasts a couple of minutes. This procedure, and variations of it, has been used for over 25 years across a wide
range of ages, populations, and experimental contexts (e.g., Carvajal & Iglesias, 1997; Ellsworth, Muir, & Hains, 1993; Gusella, Muir, & Tronick, 1988; Nadel, Croue, Mattlinger, Canet, Hudelop, & Lecuyer, 2000; Pelaez-Nogueas, Field, Hossain, & Pickens, 1996; Tronick, Als, Adamson, Wise, & Brazelton, 1978). Research reveals that infants as young as 2 months of age react to the adults’ unresponsiveness during the still-face period with decreased visual attention and positive affect, and increased self-comforting behaviors (e.g., Lamb, Morisson, & Malkin, 1987; Murray & Trevarthen, 1985; Tronick et al., 1978). In general, results show that human infants have rudimentary expectations about the nature of face-to-face interactions—they expect people to remain responsive and reciprocate in interpersonal interactions (e.g., Ellsworth et al., 1993; Muir & Hains, 1993; Rochat & Striano, 1999).

While the still-face paradigm is robust in nature, it is limited in some ways. In particular, the still-face period constitutes a dramatic and abrupt change from normal ongoing caregiver–infant interaction. During the still-face period, the infant loses maternal social contingency and reciprocity, display of maternal emotions conveyed by vocal and facial information, and often maternal touch (see Muir, 2002; Muir & Hains, 1999). To examine the separate roles of these factors in producing the still-face effect, researchers have developed a TV interaction procedure whereby mothers and infants interact over closed-circuit TV monitors (e.g., Gusella et al., 1988; Hains & Muir, 1996; Muir & Hains, 1993; Murray & Trevarthen, 1985; see also Nadel & Tremblay-Leveau, 1999). This paradigm allows only particular aspects of the interaction to be manipulated while holding other aspects constant.

Using the TV interaction procedure, Gusella et al. (1988) conducted a systematic investigation into the separate contributions of maternal facial and vocal expressions in producing the still-face effect. Six-month-old infants interacted with their mothers via closed-circuit color TVs in one of four conditions. Each condition started and ended with a 90-s normal face-to-face interaction period between mother and infant. These two normal periods were separated by either a Still face/No voice period (infants saw their mother’s expressionless, silent still-face), Still face/Interactive voice period (infants saw their mother pose a still-face while hearing her contingent voice), Interactive face/No voice period (infants saw their mother’s interactive face, but did not hear her voice), or Interactive face/Interactive voice (infants saw their mother’s interactive face and heard her interactive voice). The results revealed that when the interactive face of the mother became neutral and void of expressions during the second interaction period (Still face/No voice and Still face/Interactive voice condition), infants gazed and smiled less towards her, even if the mother’s contingent, interactive voice continued into this period. For infants whose mother’s interactive face continued into the second interaction period (Interactive face/No voice and Interactive face/Interactive voice), no decline in gaze and smile was observed relative to the normal interaction period. As long as there was a still-face, regardless of whether or not it was accompanied by an audible interactive voice, infants manifested a still-face response. As argued by the authors, a loss of contingent vocal cues does not lead to a still-face effect (see also Muir & Hains, 1999).

In some sense, this result is not surprising. Clearly, the face plays an important role in early social expectations—especially in Western cultures where face-to-face contact is prevalent (see Rochat & Striano, 1999). However, in another sense it is surprising that the voice does not influence the still-face response, given that much interaction between infants and adults also involves the voice. In many cultures infants are carried on a caregiver’s back, making vocal exchanges more important. Even in Western cultures infants are often engaged in vocal interactions with caregivers that are out of sight or while being pushed in carriages. The voice also seems to play an important role in triadic contexts such as social referencing. At 12 months of age infants use vocal cues in social referencing even when not accompanied by
facial cues (e.g., Vaish & Striano, 2004; see also Mumme, Fernald, & Herrera, 1996). Thus, the question why the voice does not appear to modulate the still-face response remains.

The current study was designed to extend upon the results of Gusella et al. (1988). In particular, we assessed whether infants manifested a still-face response when visual and vocal cues were not in conflict. It is possible that infants in the Gusella et al. study were perturbed in the Still face/Interactive voice condition because of the conflicting information they received. That is, infants may have had developed expectations that motionless, neutral still-faces are generally not accompanied by contingent vocal information. Prior research shows that infants expect faces and voices to provide congruent information (see Walker-Andrews, 1997, for a review). On the other hand, infants in the Interactive face/No voice condition might have been less disturbed given that they might have had the experience of observing people talking without necessarily hearing them (e.g., when people whisper or when people are watched from a distance). In the current study we explored these alternatives in a live interaction paradigm.

During mother–child interactions, infants received either contingent facial and vocal information (Face plus Voice), only contingent facial information (Face only), or only contingent vocal information (Voice only). The Voice only condition in the present study differed from the Still face/Interactive voice in the Gusella et al. (1988) study in that the mother could be heard but not seen, and thus, no conflicting information was provided. Based on prior research suggesting that the voice is an important aspect of interaction (Mumme et al., 1996; Vaish & Striano, 2004), we expected that infants would manifest a still-face response in all experimental conditions.

2. Method

2.1. Participants

Thirty-six full-term infants participated in the study (M = 4 months 3 days, S.D. = 10.79 days, range = 3 months 10 days to 4 months 23 days; 21 males and 15 females). An additional 13 babies were tested but not included in the study due to fussiness (n = 10), distractedness (n = 1), and experimenter error (n = 2). Participants were recruited by telephone from a database consisting of a list of names of infants whose caregivers had volunteered to participate in studies of child development. All infants were full-term and healthy, and cared for at home primarily by their biological parents. Infants were White, living in the east of Germany, and were from middle-class families. Infants received a toy for their participation.

2.2. Procedure

Testing took place in a 3 m × 4.5 m room with white walls and curtains which prevented any visual distraction. Infants were seated in a commercial infant seat. Mothers sat either 0.5 m in front of the infants at eye level or stood behind the infant and out of his/her view depending on the condition. Infants and mothers engaged in a 180-s interaction episode that was divided into three consecutive 60-s periods. As infants became available, they were randomly assigned to one of three conditions. Infants in the Face plus Voice condition engaged first in a normal 60-s face-to-face interaction with their mothers who sat in front of them (P1). This was followed by a 60-s still-face period in which the mother stopped interacting with the infant and posed a stationary, silent, still-face with neutral expression (SF). Immediately thereafter, the mother resumed a normal face-to-face interaction with the infant for the last 60-s interaction period
The three interaction periods in the Face only condition were identical to the ones in the Face plus Voice condition except that caregivers mimicked natural interaction without emitting any audible speech sounds during the first (P1) and last (P2) 60-s interaction period. In the Voice only condition, the mother stood behind the infant and a curtain so as not to be within his/her visual field, but within audible distance to the infant. Mothers saw their infants on a TV monitor to ensure contingent vocal interaction. In all other regards, the three interaction episodes were identical to the ones in the Face plus Voice condition. Thus, during P1 and P2, the infant saw and heard the mother in the Face plus Voice condition, only saw the mother in the Face only, and only heard the mother in the Voice only condition. The SF period was identical in the Face plus Voice and the Face only condition (i.e., still-face, no interactive voice), whereas in the Voice only condition the infant did not see or hear the mother (i.e., no still-face, no interactive voice).

A research assistant, who was out of view for infant and mother, timed the interaction and verbally signaled to the mother the beginning of each interaction period. Mothers were instructed to look at their infants during all interaction episodes. Mothers were allowed to use their hands to engage their child’s attention during P1 and P2 (Face plus Voice, Face only), but not to touch their infants. Infant–caregiver interactions were video-taped for later coding.

2.3. Coding

Interaction episodes were scored from video tapes by a trained coder blind to the hypotheses of the study. For each 60-s interaction period the duration of several infant behaviors were measured and converted into percentage durations. The dependent variables were operationally defined as follows:

- **Gaze (Face plus Voice, Face only):** Any looks to mother’s face.
- **Smile:** Cheeks raised and at least one corner of the mouth turned up.
- **Positive Vocalization:** Any vocalization accompanied by positive or neutral affect.
- **Negative Vocalization:** Any whimpering or insistent grunting, but excluding crying.

To assess intercoder reliability, a second, naïve coder scored a random 20% of the interaction sessions. The agreement between the two coders was Cohen’s kappa .87 for gaze, .85 for smile, .73 for positive vocalization, and .93 for negative vocalization. The percent agreement for all measures ranged from 94.5 to 99.8%.

3. Results

Preliminary results revealed no gender effects for any behavior or condition. Thus, the data were collapsed across this variable in subsequent analyses. Fig. 1 depicts the mean percent duration of gazing as a function of condition (Face plus Voice, Face only) and interaction period (P1, SF, P2). The still-face effect for Gaze was analyzed with one-way analyses of variance (ANOVA) in which interaction periods (P1, SF, P2) were treated as repeated measures. Separate repeated measures ANOVAs were conducted for the Face plus Voice and the Face only condition. The results revealed significant quadratic trends for both conditions (Face plus Voice: F(1, 11) = 9.55, p < .05; Face only: F(1, 11) = 17.67, p < .01) indicating that mean percentage duration of gazing at the mother’s face during the still-face period was significantly lower than in the preceding P1 and proceeding P2 interaction periods. Furthermore, a condition (Face plus
Fig. 1. Percent of gazing and standard error as a function of condition and interaction period.

Voice, Face only) × period (P1, SF, P2) mixed design ANOVA revealed no interaction effect (F[2, 44] = .65, p > .05), suggesting that infants’ mean looking duration in the two conditions and three interaction periods did not significantly differ from each other.

Whether Smiling decreased during the still-face period was analyzed with one-way repeated ANOVAs for the Face plus Voice, Face only, and the Voice only conditions (see Fig. 2). Only the Face only condition yielded a significant quadratic trend, F[1, 11] = 11.33, p < .05, indicating that infants smiled significantly less during the still-face period than during the P1 and P2 periods. Not finding a still-face effect in the Face plus Voice condition was surprising and perhaps due to low power. A condition (Face plus Voice, Face only, Voice only) × period (P1, SF, P2) mixed design ANOVA yielded a condition main effect, F[2, 33] = 6.72, p < .01. Least significant difference (LSD) pair-wise comparisons indicated that infants in the Voice only condition smiled significantly less than infants in the Face only condition (p < .01) and marginally less than the infants in the Face plus Voice condition (p = .051). The period main effect also reached significance, F[2, 66] = 5.50, p < .01. Across conditions, infants smiled significantly less in the still-face period than in the preceding and proceeding normal interaction period, F[1, 33] = 9.90, p < .01. Moreover, the interaction effect was significant, F[4, 66] = 4.07, p < .01, indicating that the mean duration of infant smiling during different interaction periods depended on condition. LSD pair-wise comparisons for each interaction period revealed that during P1 infants mean duration of smile was significantly higher in the Face only than in the Face plus Voice and Voice only condition. During P2, infants in the Voice only condition smiled significantly less long than the infants in the Face plus Voice and Face only condition. No other comparison reached significance.
Analogous analyses for Positive Vocalization and Negative Vocalization yielded no significant main or interaction effects (all \( p > .05 \)), suggesting that infants’ mean duration engaged in positive and negative vocalization did not vary as a function of condition or interaction period.

4. Discussion

We assessed the contribution of contingent facial and vocal information in producing the still-face effect in 4-month-old infants. We utilized a paradigm whereby infants received facial and vocal stimulation during a normal mother–infant interaction (Face plus Voice), received facial but no vocal stimulation (Face only), or received only vocal stimulation when the mother interacted with the infant from behind a curtain (Voice only). In contrast to prior studies (Gusella et al., 1988) there was no conflict between facial and vocal information. Based on research showing that infants use mothers’ positive vocal cues in social referencing, even when these are not accompanied by facial cues (e.g., Mumme et al., 1996; Vaish & Striano, 2004) we expected that infants would manifest a still-face response in the Voice only condition. This hypothesis was not supported. Rather, the current findings replicate the original findings of Gusella et al. in that the voice did not contribute to a still-face effect. These findings provide support that the still-face effect is just that—it is due to a still-face and not to a still-voice.

This finding leads to the question of why vocal information does not contribute to the still-face effect in 4-month-olds in dyadic face-to-face situations, but seems to be fundamental to social referencing by the end of the first year (e.g., Mumme et al., 1996; Vaish & Striano, 2004). The face provides key information about the referential nature of others’ behaviors and emotions and assists humans in establishing the meaning of social cues (see Moses, Baldwin, Rosicky, & Tidball, 2001; Striano & Rochat, 2000).
sensitivity to eye contact might be one explanation for the lack of still-face effect in the Voice only condition. That is, because infants did not have eye contact with their mothers in this condition, they had no access to the emotional cues conveyed by her face. However, the eye contact interpretation cannot be the entire explanation for at least two reasons. First, from birth, infants are sensitive to eye contact (e.g., Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Massaccesi, & Simion, 2002). However, there is no reported still-face effect at birth, suggesting that social experience may be necessary for human infants to develop social expectations. Second, infants between 3 and 9 months of age manifest a robust still-face effect even if the adult is not looking at them (see Delgado, Messinger, & Yale, 2002; Striano, 2004).

Infants in the Face plus Voice and Face only condition had contingent facial cues disrupted during the SF period—facial cues to which infants in the Voice only condition had no access. A particular visual cue alone or in concert with vocal cues might be essential for the still-face effect in 4-month-olds. This idea is in line with theories suggesting that the recognition and discrimination of emotional expressions is enhanced with multimodally presented information (face and voice) compared to information presented from only one modality (face or voice) (see Walker-Andrews, 1997; also Walker-Andrews & Lennon, 1991). Furthermore, it is possible that smiling contributed to the still-face response. Infants in the Face plus Voice and Face only conditions saw their mother’s smile abruptly disappear from the P1 to the still-face period. The results showed that infants’ attention to the mother’s face declined significantly between these two interaction periods accompanied by a drop in positive affect. Infants in the Voice only condition exhibited relatively little positive affect in all three interaction periods, indicating that only hearing a mother’s positive interactive voice is not enough to elicit smiles in infants. Thus, the still-face reaction might reflect a sensitivity to an interaction partner’s smiling expression. This hypothesis is corroborated by research that found that infants look reliably longer to smiling than to neutral expressions (e.g., Kuchuk, Vibbert, & Bornstein, 1986; Striano, Brennan, & Vanman, 2002). However, research has also shown that maintaining a positive facial expression during the still-face period does not play a strong role in modulating infants’ still-face reaction (e.g., D’Entremont & Muir, 1997; Rochat, Striano, & Blatt, 2002). These findings have led to the conclusion that the loss of interpersonal contingency is a major contributor to the still-face. If this is the case, the current study shows it is a loss of facial contingency and not vocal contingency.

The observed difference in still-face effect between the Face plus Voice, Face only and the Voice only condition might also lie in infants’ sensitivity to facial movement. In both the Face plus Voice and Face only condition there was a loss of the interaction partners’ facial movement during the still-face period. Infants might simply react to the sudden absence of facial movement with decreased visual attention and positive affect. Research has shown that infants are sensitive to movement and visually prefer a moving over a static stimulus (e.g., Dannemiller & Freedland, 1989; Hicks & Richards, 1998; Lewis, Maurer, Burkhanpurkar, & Anvari, 1996). Furthermore, Ellisworth et al. (1993) found that 3- and 6-month-olds exhibited the same still-face effect for gaze whether a person or an interactive object (a hand puppet whose internal features moved) became still-faced. The authors argue that “...infants’ visual attention may have been driven primarily by movement...” (p. 68).

While we alluded to possible explanations for the observed differences in the still-face response in 4-month-olds, the question of why infants failed to exhibit a still-face effect in the Voice only condition when, just a few months later, they are adept at using vocal cues to guide their own behavior (Mumme et al., 1996; Vaish & Striano, 2004) remains. Future studies are needed to address these questions. For example, there is scant research relating developing dyadic skills such as social expectations in the still-face paradigm to triadic behaviors that require infants to understand someone’s expression as intended.
for the self but referring to something else. One area for future research could be to assess infants’ behavior in dyadic and triadic contexts to determine if and when there are strong relations. Another area of future research is to test infants’ social expectations as a function of development. This approach would establish if learning and experience play a role in infants’ understanding of the relevance of vocal cues that are not accompanied by faces. A developmental approach is also critical in teasing apart the relation between maturation and experience. For example, if 12-month-old infants rely upon vocal cues in social referencing but younger infants do not, can this development be better explained by the emergence of locomotion or developing action systems, rather than more general experience? With developing action systems such as independent locomotion, the need for information changes as does the contexts in which it is provided. Studies are needed to determine how action systems and the need for information control the channels of communication that are selected and used by infants. It would be useful to know for instance, if the significance of facial cues declines over the first year and if the importance of vocal cues increases as a function of locomotion and the need to social reference and use emotional cues when far from the caregiver (see Campos, Anderson, Barbu-Roth, Hubbard, Hertenstein, & Witherington, 2000).

The current study tested only 4-month-olds and was limited in sample size. In addition to testing more infants, it should be assessed if older infants are more inclined to respond to a loss of contingent vocal cues in dyadic contexts, and if these are the same infants who use vocal cues in social referencing (triadic) tasks. The voice is an integral part of social skills (see Fernald, 2001) and in establishing the meaning of expressions among adults (Schirmer, Kotz, & Friederici, 2002), but it is not a determinant of the still-face response in 4-month-olds. The key is to establish in which contexts vocal and facial cues are perceived as relevant and the developmental course of these sensitivities.

Acknowledgments

We would like to thank the parents and the infants who participated in this study, Caterina Böttcher and Kerstin Träger for collecting and coding data, and the Universitätsfrauenklinik, Leipzig for support and assistance in subject recruitment.

References

