Five- and Eight-Month-Old Infants Recognize Their Faces and Voices as Familiar and Social Stimuli

Maria Legerstee, Diane Anderson, and Alliza Schaffer

Five- and 8-month-old infants were presented with silent moving and static video images of self, peer, and doll and with sounds of self, peer, and nonsocial objects. In the visual conditions, infants at both ages showed a significant looking preference to peer over self when the faces were moving. When the faces were static, older infants showed the same significant discriminations, but the younger infants showed a significant looking preference for their own faces over peer and doll. These data suggest that recognition of one’s own image develops through experience with dynamic facial stimulation during the first 8 months of life. In the auditory conditions, infants at both ages showed significant looking preferences for sounds of peer over self or nonsocial objects. In general, infants of both ages smiled and produced more vocalizations to social faces and social sounds than to nonsocial faces and nonsocial sounds. Thus, at 5 months infants treat their faces and voices as familiar and social stimuli. The findings forge important links among studies of self-perception, self-recognition, and social knowledge.

INTRODUCTION

From birth, infants actively use their visual and auditory systems to acquire information about their surroundings and themselves. Infants are visually attracted to movement, contour, contrast, certain levels of complexity, and curvature (Banks & Salapatek, 1983; Haith, 1966; Salapatek & Kessen, 1966). Newborns seem to prefer these parameters when they are arranged in face-like rather than in abstract ways, because they track a human schematic face significantly more than a scrambled face or a blank face (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991; Johnson & Morton, 1991). By 2 months, infants discriminate between the faces of mothers and female strangers (Barrera & Maurer, 1981), and by 4 months, they distinguish among the faces of a man, woman, and baby (Fagan, 1972). The results suggest that rather than having to construct a notion of a human face out of the various physical parameters mentioned above, infants may be born with some kind of template for “faceness” (e.g., primal specification of some structural characteristics of the human face called “CONSPEC” by Johnson & Morton, 1991, p. 105). Newborn infants spend more time looking at their mothers’ faces than at strangers’ faces (Bushnell, Sai, & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984). This demonstrates that they are able to recognize their mothers’ faces as familiar and suggests that this activity cannot be regulated by sensory information alone (e.g., “CONSPEC,” Johnson & Morton, 1991); it must also involve information stored in memory.

Auditory perception also appears to be well developed in newborns. Unlike the visual system, the auditory system is stimulated in utero (DeCasper, Lecanuet, Busnel, Granier-Defere, & Naugeais, 1994). Newborns attend preferentially to human speech over other sounds (DeCasper & Fifer, 1980). At 1 month, they make fine distinctions among speech sounds (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) and discriminate between linguistic contrasts that are not available in their mother tongue (Trehub, 1976). By 4 months, infants can reliably hear sounds and determine the direction the sound originates from (Muir, Clifton, & Clarkson, 1989). Newborns are able to represent human sounds; they recognize the voices of their mothers over those of female strangers (DeCasper & Fifer, 1980) and retain information about syllables (Jusczyk, Kennedy, & Jusczyk, 1995).

Not only do infants represent others’ faces and voices during the first 6 months of life, they seem to be able to recognize their own faces and voices as familiar stimuli (Bahrick, 1995; Bahrick, Moss, & Fadil, in press; Martin & Clark, 1982). Such findings are of interest because they shed light on the beginnings of self-recognition and social development in infants. In a review of the developmental perspectives of the self-system, Harter (1983) writes that throughout history the face seems to have been regarded as the primary representation of the self. She notes that the face, as an embodiment of the self, can be found in such metaphorical expressions as “I won’t be able to face myself in the mirror in the morning” (p. 218). Gallup (1982) demonstrated that the
recognition of one's own physical features is crucial for the development of an awareness of self. In a variety of studies, he found that after 3 days of experience with mirrors, chimps would change from making threatening gestures to grooming gestures, indicating that they gradually recognized themselves in the mirror. Behaviors of human infants to their mirror images in early self-recognition studies by Papousek and Papousek (1974) and Lewis and Brooks-Gunn (1979) showed that infants between 5 and 8 months could discriminate between live video images of themselves and recorded images of peers. The authors suggested that self-perception was based on the infants' recognition of differential contingencies between the visual stimulation of the video images and the proprioceptive feedback from their own body motion. Field (1979) also found that 3-month-old infants tended to spend more time looking at mirror images of self than at a noncontingent presentation of peers, although they smiled and vocalized more toward peers. It could be argued that infants responded differentially because there were differences between eye contact and eye motion. However, Bahrick and Watson (1985) separated featural information from contingency feedback by presenting infants with video images of their own legs and those of other infants. Five-month-old infants discriminated between the live film of their own leg movements and that of another infant's movements, and they discriminated between live and prerecorded video images of their own legs in a condition where the infants' direct views of their own bodies were occluded. In both conditions, infants looked longer at the noncontingent view. The authors suggested that infants were able to make these distinctions because they were sensitive to visual and proprioceptive contingencies of their own (familiar) movements.

If infants are able to recognize their own legs because in real life they have associated proprioceptive movements with visual images of their legs, then this process should allow infants to become familiar with other aspects of self, such as their faces and voices. In an intriguing set of investigations of the facial self, Bahrick et al. (in press) presented 5- and 8-month-olds with previously recorded faces of self and peer in conditions where the images were either moving or frozen. The faces were presented in four positions: (1) straight ahead, (2) 45° to the right or left, and (3) looking up. For uniformity, infants' fronts were covered with yellow bibs, and to control for emotional expressions and movement a yoked-control design was used so that the face of each infant served as the face of the peer for the next infant. Infants were presented with four 30 s trials in the moving condition and four 15 s trials in the frozen condition. Infants preferred to look at the moving face of a peer over self (novelty preference). A replication of this study with 2- and 3-month-old infants showed no discrimination between the faces by 2 months, but an emerging preference for the faces of peers by the second block of trials in the 3-month-old infants. The authors suggested that infants' abilities to distinguish their own faces from those of peers emerged through prior visual experience with their own faces in the mirror, where proprioceptive contingencies identify the face of the self and differentiate it from other faces. Other researchers have shown that infants also appear sensitive to proprioceptive movements of their hands and arms (Schmuckler, 1996) and their bodies (Rochat, 1995).

Aside from seeing and feeling themselves move, infants from birth experience contrasting auditory events that should inform them about their own sounds. When infants vocalize, the sounds they hear are paired with kinesthetic and proprioceptive feedback. Sounds produced by other people or objects do not provide infants with such intermodal information. However, there are few empirical data suggesting that young infants perceive their voices as familiar. Martin and Clark (1982) conducted a study to investigate the often reported tendency of newborns to cry as a result of hearing other infants cry. They compared the responses of day-old infants to prerecorded sounds of their own cries as well as the cries of other newborns. Their results showed that infants discriminated their own cries from those of other infants; babies that were calm at the commencement of the experiment vocalized considerably more when hearing the cries of other infants, but less when hearing their own cries. It has been suggested that these findings provide evidence for an auditory specification of self in newborns (Butterworth, 1990).

The above evidence seems to suggest that infants perceive, through various forms of proprioception, their own vocal, facial, and bodily movements and that they gain an understanding of themselves as physical agents through observing the effects of these actions (Tommasello, 1995). Thus, situations that seem to give rise to the physical or ecological self (cf. Neisser, 1995) are continuously specified by visual, acoustic, kinesthetic, and vestibular information. It has been suggested, however, that a sense of self as a social entity must be gained through interaction with the social environment (Legerstee, 1997a, in press; Neisser, 1995; Reed, 1995). From birth, infants demonstrate their abilities and motivation to interact with
people (Neisser, 1995; Stern, 1995; Trevarthen, 1979). The infant develops a mode of self-knowledge from earliest infancy, through interactions that are reciprocally controlled by human signals such as smiles and vocalizations (Neisser, 1995). Through these interactions, infants soon demonstrate that they have developed certain expectations. When 2- to 3-month-olds are presented alternately with responsive and passive people (mother and female stranger) and objects (dolls with schematic faces, partial faces, or no facial features) in naturally interactive paradigms, infants will smile and vocalize more to responsive people than to “interactive” objects, whether the objects have facial features or not (Ellsworth, Muir, & Hains, 1993; Legerstee, 1997b; Legerstee, Corter, & Kienapple, 1990; Legerstee, Pomerleau, Malcuit, & Feider, 1987). The results cannot be related to a difference in movement pattern between the social and nonsocial stimuli, because infants respond differentially also when the stimuli remain immobile. They show distress to unresponsive people but not to static objects. It has been proposed that infants expect people to communicate reciprocally with them in face-to-face interactions, each partner working actively to sustain and regulate the interaction (Stern, 1995; Trevarthen, 1979; Tronick, Ricks, & Cohn, 1982). Gibson (1995, p. 12) proposes that information about the self is first obtained in the discovery of control; it is the expectation of intentions in others that leads to the development of self as a social agent. Thus, it would appear that, whereas infants learn about the physical self through interacting with physical things, they learn about the social self through interacting with social objects. It has already been established that, by 2–3 months, infants respond differentially to people and things; they treat people as social objects, smiling, vocalizing, and imitating their actions, but objects are treated as toys to be looked at and manipulated (for reviews, see Legerstee, 1992, 1997a, in press; Spelke, Phillips, & Woodward, 1995). However, the claim that infants may perceive themselves as similar to other social entities and different from nonsocial objects does not appear to have been tested.

The particular aim of the current research was to determine whether infants (1) discriminate their own faces from those of peers and dolls, (2) discriminate their own voices from those of peers and nonsocial sounds, and (3) respond differentially to these social and nonsocial objects. Consequently, 5- and 8-month-olds were presented with their own silent moving and static faces, the faces of peers, and dolls; they also listened to their own voices, those of peers, and to various nonsocial sounds. The infants’ looking, vocalizations, and smiles were coded in each condition. It was hypothesized that if infants perceive their own faces and voices as familiar, they would look longer toward novel faces and voices than toward familiar faces and voices (Bahrick et al., in press; Johnson & Morton, 1991). Furthermore, if infants differentiated between the social and nonsocial faces and voices, then they would produce more smiles and vocalizations when presented with social stimuli than when faced with nonsocial stimuli (Ellsworth et al., 1993; Legerstee, 1997b; Legerstee et al., 1987, 1990).

**EXPERIMENT 1**

In this study, infants were presented with prerecorded video films of their own faces, those of peers, and those of dolls in both static and moving conditions. Studies have shown that movement facilitates discrimination between faces of peer and self in 3- and 5-month-olds (Bahrick et al., in press) and between schematic faces of people and dolls in 5-month-olds (Johnson et al., 1991). Infants were further presented with previously recorded sounds of self and peer, and with nonsocial sounds. To determine whether infants discriminated among the social and nonsocial faces and sounds, their looking, smiles, and vocalizations were recorded. The index of preference/discrimination is the difference in the average length of infants’ overall looking time, smiling, and vocalizing to the three different kinds of stimuli in the static and moving visual conditions, and in the auditory conditions.

**General Method**

**Participants**

A total of 51 5-month-olds and 30 8-month-olds were recruited from the North York General Hospital. Of those infants, nine 5-month-olds and six 8-month-olds were excluded because of experimenter error or because the infant did not achieve the necessary quiet, alert state during the session. The remaining 42 5-month-old infants (18 girls) and 24 8-month-olds (18 girls) were healthy term babies, ranging in weight from 2,240 to 4,928 g ($M = 3,332, SD = 59$) at birth; they were tested at 4–5 months ($M = 5.1, SD = .13$) or 7–8 months ($M = 8.07, SD = .14$).

There were 14 5-month-olds and eight 8-month-olds in each of the three conditions: (1) visual static, (2) visual moving, and (3) auditory. We chose to observe only eight 8-month-old babies for each of the
three conditions because research has shown that, at this age, infants discriminate between the faces of self and peer (Bahrick et al., in press) as well as between people and objects (Ellsworth et al., 1993; Legerstee, 1994, 1997b; Legerstee et al., 1990, 1987).

Procedure

The experiment was conducted in the laboratory and transpired over two visits. Prior to commencement of the first phase, the parents were asked to fill out a questionnaire about the amount of time their infants saw themselves in the mirror and what their subsequent reactions were, and whether the infants possessed toys with facial features. All infants saw themselves in the mirror at least once a day during caretaking activities. Most mothers (95%) reported that the infants looked at the caregivers before looking at themselves and then began to smile or coo at themselves. All infants had some type of a soft doll or animal with facial features that made sounds when handled.

Visit 1. After completion of the infant questionnaire, parent and infant were taken to the assessment room where the infant was clothed in a yellow robe, leaving only the face and hands uncovered. The yellow robes created uniformity and eliminated any potential visual distractions that could affect the infants' responses during the second phase. They also controlled for color and brightness of the infants' clothes. Infants were placed in a blue foam-filled armchair especially designed for young babies. Their heads were supported on both sides but were free to rotate. The chair was mounted on top of a black table measuring 98 cm x 73 cm in width and 24 cm in height. During this visit, the mother was seated approximately 30 cm away from the infant's face. Infant and mother were enclosed by curtains on four sides. Behind the mother, and between two of the curtains, a video camera with a zoom lens recorded the infant's vocal and behavioral responses during the experiment. To enhance visibility of the faces on the demonstration tape, the camera operator used the zoom lens to get maximal facial exposure of infants or doll and minimal background (chair and frock) on the video monitor. Once the infant attained an alert, content state (stage 4; Wolff, 1966), the mother was given a toy and asked to interact with her baby as she naturally would at home, but not to speak to the baby, for approximately 5 min. Of these 5 min tapes, a 60 s demonstration tape was prepared for use during the second phase. By forwarding through the 5 min interaction tape the experimenter edited a 2 s section of continuous vocalizations of the baby, repeated this section three times, and finished it with a 1 s pause during which the baby was seen (in the visual conditions) but did not vocalize. After this 7 s sequence the experimenter continued to search for the next vocalizations on the 5 min tape and edited a subsequent sequence of 2 s vocalizations, repeated the sequence three times, and finished it with a 1 s pause. These sequences were repeated until the duration of the demonstration tape was 60 s. Thus, the demonstration tapes showed babies who moved their faces and vocalized in a burst-and-pause pattern (three successive bursts of vocalizations of a total of approximately 6 s and a 1 s pause during which the baby did not vocalize, repeated for a total of 60 s). We chose a burst-and-pause pattern because it has been shown that a paradigm during which the "infant perceives a human [adult] acting, then stopping, acting, then stopping, may motivate the infant to action, rather than to mere visual fixation" (Meltzoff & Moore, 1983, p. 290).

Visit 2. The second visit for all infants transpired between 4 and 6 days after the first appointment. Prior to the second visit, infants had been assigned to one of the three possible conditions: (1) visual moving, (2) visual static, or (3) auditory.

In the interim between visits, a 60 s tape of the visual and sounding nonsocial object had been constructed to match with the target infant and a peer. To provide a close match with respect to degree of visual complexity, the experimenter moved the object and yoked it to the facial movements of the target infant. The camera operator filmed the object for the demonstration tape by focusing the zoom lens of the camera on the doll face so that it filled the entire screen to keep the background of the social object equivalent to that of the nonsocial object. Each object was made of a 12.7 cm foam rubber ball (size of baby’s head) attached to a stick and clothed in a yellow robe (see Figure 1). The object had removable facial features that were arranged to represent a normal face. Hair (if necessary, because not all babies had hair) was matched to the target infant and peer with yellow, various shades of brown, or black yarn. More than one object was used to provide a closer match to the target infant and the peer with respect to hair and skin color and to make sure that no idiosyncratic feature of a single object could influence the discrimination of the social and nonsocial stimuli.

More than one nonsocial sound was also used. The sound was made either by bells or by a synthesizer (determined randomly, so that half of the babies received the bell sound and half received the synthesizer). For the bell sound, bells (three differently sized shaker bells) were shaken and yoked to the vocaliza-
tions of the infant on the auditory demonstration tape (three 2 s bursts; dB = 64–67, and a 1 s pause, repeated for a total of 60 s). When the synthesizer was used, a modified version of jingle bells up and down the scales was played (one octave above middle C; dB = 64–67) in a pattern of three 2 s sounds (three to five notes in a single pitch repeated three times) and a 1 s pause, and followed by more rapid runs of four to eight notes up and down the scales for a total of 60 s.

Preparation of the demonstration tapes ensured that the infants were provided with relatively equivalent stimulation concerning complexity, movement, rhythm, and intensity of the image and sound of the peer, object, and self (all stimuli moved in a 6 s burst, 1 s pause pattern). Target infants and peers were further matched for smiles, hair, and skin color. Because every infant served as the peer for another infant, other physical characteristics were distributed unsystematically.

Infants were presented alternately with the two social and the one nonsocial stimuli during each session. Pilot work had shown that the 1 min visual or auditory stimulus presentation was long enough to assess discrimination. No more than three 1 min displays were presented, because the presentation of many different stimuli might confuse infants and decrease their level of affect (Ellsworth et al., 1993). A latin square was used to determine the order of presentation of the between-subjects variable conditions: (1) moving (M), (2) static (S), and (3) auditory (A). The orders MSA and SAM each occurred five times, and the order ASM occurred four times for the 5-month-olds. The orders AMS and MSA each occurred three times, and the order SAM occurred twice for the 8-month-olds. Order of presentation of the stimulus displays (Self, Peer, and Object) was randomized across infants and conditions in such a way that each of the possible six orders occurred seven times across the 42 5-month-olds and four times each across the 24 8-month-olds.

During the static condition, the infants saw a frozen image of self, peer, or doll with sound off; during the moving condition, the infants saw a silent but...
moving visual image of self, peer, or doll, and during the auditory condition, the infants heard their own vocalizations, those of peers, or nonsocial sounds (either bells or synthesizer) but saw no images. Infants were placed in the baby seat on top of the table with a 45 cm color television located at a 75 cm distance. The infant and television were enclosed in a confined area measuring 152 cm × 152 cm and surrounded by curtains. Infants were given a few minutes to become familiar with the experimental setup until they were in a relaxed and attentive state (Wolff, 1966).

Visual Conditions

Method

Although each infant was presented with all three conditions, the procedures were somewhat different for the visual conditions and the auditory condition. The data are, therefore, presented separately.

Procedures. During the visual conditions, infants had to lift their heads slightly to look at the display; a camera that was mounted just above the television screen captured their eyes. The experimenter first attracted the infant’s attention to the screen with a flashing red light or by tapping on the screen from behind. Once the infant looked up to the screen and in the right direction, one of the three visual stimuli was displayed on the television screen for a total of 60 s.

Measures and scoring. We hypothesized that if infants recognized their own facial images as familiar, they would look less long at their own images than at the other stimuli. In addition, if they perceived images of self and peer as social then they would produce more social behaviors to the peers and themselves than to the object. Social responsiveness was defined in terms of the number of smiles (corners of the mouth turned up) and vocalizations (cooing sounds). Crying and other sounds (coughing, burping, sneezing, and so on) were excluded.

For the entire experiment, the durations of infants’ looking and the frequencies of smiles and vocalizations were coded by people who were unaware of which stimulus had been presented. Coding of looking began when infants looked at the stimulus display and discontinued when infants looked away. Before coding began, a time generator clocked the taped segments by 0.1 s. Two coders, trained on 30% of the data, coded smiling and the onset and offset of infants’ looking while the sound was turned down, and two different people coded the vocalizations while the picture was turned down. Reliability was measured by Cohen’s kappa on seconds of agreement for looking, .81, and on frequency rates for vocalizations, .89, and smiles, .81.

Results and Discussion

The infant responses were analyzed by a three-way ANOVA in which stimulus (3: person, peer, object) was the within-subjects factor, and age (2: 5 and 8 months) and condition (2: moving and static) were the between-subjects factors. Means and standard deviations for looking, smiles, and vocalizations are reported in Table 1. When age was not significant, the data for the 5- and 8-month-old groups were collapsed and analyzed separately by a two-way ANOVA in which stimulus was the within-subjects factor and condition the between-subjects factor. All within-subjects effects are reported with Greenhouse-Geisser adjusted $p$ values. Significant interactions were analyzed by using the simple effects method. Stimulus levels were compared with single degree of freedom nonorthogonal contrast tests (Winer, 1972). The statistical package of SPSS accounts for the differences in sample size between the 5- and 8-months age groups by giving equal weight to all cells.

Looking. As Figure 2 shows, there was a significant age × condition × stimulus interaction, $F(2, 80) = 9.64, p = .001$. Subsequently, a two-way ANOVA was conducted on each age group separately. For the 5-month-old group there was a significant condition × stimulus interaction, $F(2, 52) = 19.82, p = .001$, suggesting that the pattern of responding for the two conditions was different. Within the moving condition, the contrast tests showed that infants looked reliably longer at the unfamiliar faces of the peers than at themselves, $F(1, 26) = 7.24, p = .012$, and they also looked longer at object than at self, $F(1, 26) = 6.69, p = .016$. In contrast, within the static condition contrast tests showed that infants looked significantly longer at their own still faces, $F(1, 26) = 16.85, p = .001$, than at those of peers and objects, between which there was no significant difference.

For the 8-month-old group, there was only a significant main effect of stimulus, $F(1, 28) = 27.38, p = .001$, indicating that the pattern of responding to the various stimuli was similar for the two conditions. Subsequent contrast tests showed that infants demonstrated a significant looking preference for peer over self, $F(1, 14) = 43.82, p = .001$, for object over self, $F(1, 14) = 12.43, p = .003$, and for peer over object, $F(1, 14) = 5.27, p = .037$.

Smiles. None of the age and condition effects were significant. A main effect for stimulus, $F(2, 84) = 7.36, p = .001$, and contrast tests suggested that the infants showed similar patterns of responding to people and objects at both ages and in both conditions. Infants smiled more to self and peer (between which there was no significant difference) than to the object, $F(1, 42) = 13.22, p = .001$. 
Table 1 Means and Standard Deviations of Infant Responses as a Function of Age, Visual Moving and Static Conditions, and Stimuli

<table>
<thead>
<tr>
<th>Measure and Stimulus</th>
<th>n</th>
<th>Moving</th>
<th>Static</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Looking:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td>14</td>
<td>27.8</td>
<td>14.1</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>37.6</td>
<td>10.8</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>45.8</td>
<td>10.1</td>
</tr>
<tr>
<td>8 months:</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td>18.8</td>
<td>7.6</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>36.3</td>
<td>12.9</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>28.6</td>
<td>14.1</td>
</tr>
<tr>
<td>Smiles:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 months:</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td>3.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>2.2</td>
<td>7.1</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>.6</td>
<td>1.6</td>
</tr>
<tr>
<td>8 months:</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td>1.3</td>
<td>1.5</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>.8</td>
<td>.7</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>.13</td>
<td>.3</td>
</tr>
<tr>
<td>Vocalizations:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 months:</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td>4.1</td>
<td>5.4</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>3.2</td>
<td>5.8</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>1.0</td>
<td>3.1</td>
</tr>
<tr>
<td>8 months:</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td>3.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Peer</td>
<td></td>
<td>2.7</td>
<td>2.9</td>
</tr>
<tr>
<td>Object</td>
<td></td>
<td>1.25</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Vocalizations. Again, none of the age or condition effects were significant, but a main effect for stimulus, $F(2, 84) = 8.11, p = .001$, and subsequent contrast tests showed that infants vocalized significantly more to self and peer (between which there was no significant difference) than to the object, $F(1, 42) = 27.56, p = .001$.

Summary

The results demonstrate that 5- and 8-month-old infants discriminated between the faces and the social and nonsocial stimuli. Infants of both ages looked longer at the novel video images of peer and object than at their own faces; 8-month-olds also preferred...
the novel static displays, but 5-month-olds preferred their own faces over the unfamiliar static stimuli. Both smiles and vocalizations were directed primarily toward human faces.

**Auditory Condition**

**Method**

*Procedure.* The auditory condition followed a procedure different from that of the visual conditions. Muir et al. (1989) demonstrated that by 5 months, infants can reliably localize off-centered sounds by turning their heads to the direction of the sound. Accordingly, infants faced forward, and the video screen was situated 75 cm away, rotated 90° left or right (determined randomly) from the central fixation point. The picture on the monitor was turned off and covered by a curtain. For the infants to localize the sound they had to turn their heads toward the sound and look straight into the camera that was placed above the sound source. Infants were tested in dim light to keep visual distractions to a minimum. To attract the infant’s attention an experimenter tapped on top of the television monitor. All infants turned toward the tapping sound. When the infant looked toward the tapping sound, the sound of one of the three tapes (Self, Peer, or Object) was played.

*Measures and scoring.* Looks, smiles, and vocalizations were used as measures of discrimination. Coding of looking began as soon as the infant looked toward the stimulus and terminated when the infant looked away. Again, two coders were trained on 30% of the data. They coded the smiles and the onset and offset times of looking toward the stimulus. As in the visual conditions, two different coders coded the vocalizations of the babies. The coders were unaware of our hypotheses. Reliability was analyzed by Cohen’s kappa on the agreement in seconds for looking, .82, and on frequency rates for vocalizations, .81, and smiles, .81.

**Results and Discussion**

Infant responses to the auditory stimuli were first submitted to a repeated-measures ANOVA with age (2: 2 and 8 months) as the between-subjects factor and stimulus (3: Self, Peer, and Object) as the within-subjects factor. Means and standard deviations of looking, smiles, and vocalizations are reported in Table 2. Because none of the results with age were significant, the data of the 5- and 8-month-olds were collapsed and submitted to a one-way ANOVA. Single degree of freedom nonorthogonal contrast tests were performed to assess for significant differences between the stimuli (Winer, 1972).

*Looking.* The responses to the auditory stimuli revealed a changing pattern in visual attention as a function of the three stimuli. A significant main effect of stimulus, $F(2, 42) = 32.83, p = .001$, and subsequent contrast tests revealed that infants looked longer at the sounds of peer than those of self, $F(1, 21) = 15.61, p = .001$, but they looked longer at sounds of self and peer than at nonsocial sounds, $F(1, 21) = 76.94, p = .001$.

*Smiles.* A significant main effect of stimulus, $F(2, 42) = 5.16, p = .01$, and subsequent contrast tests showed that the infants smiled reliably more when they heard their own vocalizations and those of the peer, than when they heard the nonsocial sounds, $F(1, 21) = 8.78, p = .007$.

*Vocalizations.* A main effect of stimulus, $F(2, 42) = 12.72, p = .001$, and subsequent contrast tests showed, interestingly, that all infants vocalized significantly more to their own voices than to those of peers, $F(1, 21) = 10.82, p = .003$, and they vocalized least when hearing the nonsocial sounds, $F(1, 21) = 15.07, p = .001$.

**Summary**

The results showed that 5- and 8-month-olds discriminated among the sounds of the peer, object, and self. They looked longer toward the peers’ voices than toward their own and least long at the inanimate sounds. Smiles and vocalizations occurred most frequently when infants heard human voices, with the most frequent vocalizations occurring when they listened to their own voices.

**EXPERIMENT 2**

The finding that infants produce more social responses to people than to nonsocial objects is consis-
tent with an interpretation that infants recognize faces as social stimuli (Bruner, 1973; Gelman, Durgin, & Kaufman, 1995; Legerstee, 1997a; Mandler, 1992; Spelke et al., 1995; Trevarthen, 1979). However, in the natural environment, people move their facial features whereas nonsocial objects do not; therefore, in Experiment 1, it is possible that infants smiled and vocalized at people, not because they differentiated people from objects, but because the objects moved externally rather than internally with deforming motions. The next experiment was carried out to clarify this issue.

In Experiment 2, the visual moving condition was replicated with 5-month-olds only. This time, however, the nonsocial object was constructed with moveable facial features arranged in an abstract way (see Figure 1). An abstract pattern was chosen over a facial pattern, to provide a basis for discriminating between the video images of the human face and the doll’s face.

Method

Participants. Ten infants (M = 5.3 months, SD = .19) participated in Experiment 2. Data from three additional infants were eliminated, two because of crying during the testing procedure and one because of experimenter error.

Procedure. Again the experiment occurred over two visits and was conducted in the Infancy Laboratory. Results of the questionnaires indicated that all infants saw themselves once or twice a day in the mirror. Most mothers (90%) reported that the infants looked first at their mothers and then at themselves, and that the infants smiled and vocalized most of the time at their own faces. All infants had a toy with facial features that made sounds when handled. The purpose of Visit 1 was to gather data to construct demonstration tapes for Visit 2. The procedure during Visit 1 for the creation of the 60 s demonstration tape was identical to that of Experiment 1.

The second visit transpired between 4 and 6 days after the first visit. All infants were presented with the internal moving condition of the peer, self, and object. Again a 60 s demonstration tape of the moving nonsocial object was constructed. This time the object was a terry cloth, multicolored puppet (see Figure 1). The experimenter put her hand inside the puppet, while a yellow robe was draped around her arm. She moved the internal features of the puppet so that the deforming motions of the puppet were yoked to the facial movements of the target infant on the social demonstration tape, in a 6 s burst and 1 s pause pattern for a total of 60 s. Again the camera operator focused the zoom lens so that the puppet’s face filled the entire screen to keep the background of the social and nonsocial faces equivalent.

The procedure of the internal moving conditions was identical to that of the visual moving condition in Experiment 1. Order of presentation of the visual displays, Self (S), Peer (P), and Object (O) was randomized, so that the six possible combinations were administered to the first six babies, and the combinations PSO, SPO, POS, and OSP to the remaining four infants. Again each infant served once as target infant and once as peer.

Measures and scoring. Looks, smiles, and vocalizations were used as measures of discrimination. Coding was performed as in Experiment 1, and reliability was coded for 30% of the data by coders who were unaware which stimulus had been presented. Reliability was measured by Cohen’s kappa on agreement in seconds for looking, .81, and on frequency rates for vocalizations, .89, and smiles, .83.

Results and Discussion

Infant responses to the internal moving stimulus were submitted to a one-way ANOVA with stimulus (3: Self, Peer, and Object) as the within-subjects factor. Contrast tests were performed to assess for significant differences among the stimuli. Means and standard deviations for looking, smiles, and vocalizations are reported in Table 3.

Table 3 Means and Standard Deviations of Infant Responses to Self, Peer, and Object in the Internal Movement Condition

<table>
<thead>
<tr>
<th>Measure and Stimulus</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Looking, 5 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td>10</td>
<td>23.0</td>
<td>13.7</td>
</tr>
<tr>
<td>Peer</td>
<td>31.6</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td>Object</td>
<td>45.3</td>
<td>16.3</td>
<td></td>
</tr>
<tr>
<td>Smiles, 5 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td>3.6</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>Peer</td>
<td>3.7</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>Object</td>
<td>1.0</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>Vocalizations, 5 months:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td>5.4</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Peer</td>
<td>3.1</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Object</td>
<td>1.4</td>
<td>2.6</td>
<td></td>
</tr>
</tbody>
</table>
indicated that infants smiled significantly less at the object than at faces of self and peer, between which there was no significant difference, F(1, 9) = 9.35, p = .014.

Vocalizations. Again, a main effect of stimulus, F(2, 18) = 7.50, p = .013, and subsequent contrast tests revealed that this time the difference in the vocalizations to peer and self was significant. Infants vocalized more at self than at the peer, F(1, 9) = 12.49, p = .006, and more at the peer than at the object, F(1, 9) = 6.27, p = .034.

Summary

The results of Experiment 2 indicated that despite the internally moving features of the nonsocial stimulus, infants again distinguished between the faces of people and objects with their looks, smiles, and vocalizations. They looked longer at the novel face of the puppet than at the peers' faces, and they looked least long at their own faces. As in Experiment 1, the infants reserved their smiles and vocalizations primarily for people, with the most vocalizations occurring when infants were watching their own faces.

The above findings make it unlikely that the results of Experiment 1 could have been due to the difference in internal motion only. Recall that in Experiment 1 the objects had facial features arranged canonically but they were static; in Experiment 2, the object had features that moved relative to each other, but the features were scrambled. This means that the infants could have discriminated human faces from objects because of lack of internal movement in Experiment 1 and lack of facial configuration in Experiment 2. This suggests that infants use multiple bases for discriminating between people and objects.

GENERAL DISCUSSION

These studies show that infants as young as 5 months discriminate the moving video images of themselves from those of peers and dolls, and their own sounds from those of peers and nonsocial objects. The early discrimination of such complex visual and auditory patterns is quite impressive and suggests that infants have a well-developed ability to process visual and auditory information. The infants not only discriminated among the stimuli, but they appeared to recognize their own faces and voices as familiar as indexed by a novelty preference for the faces and sounds of the peer and object. They differentiated among the social and nonsocial faces and sounds in another way by smiling and vocalizing more to the former than to the latter.

During the moving condition, the 5-month-olds looked less long at their own faces than at those of the peers, a finding consistent with Bahrick et al. (in press), and they looked longer at the objects than at the peers. In the static condition, however, the 5-month-olds looked longer at their own faces than those of peers and objects. Johnson et al. (1991, Experiment 2) also found that 5-month-olds did not discriminate with their looking among static schematic faces and other patterns, but if the internal features were moving, infants preferred human faces over the other stimuli.

The finding that both 5- and 8-month-olds looked less long at their own faces (familiarity effect) in the moving condition but not in the static condition suggests that movement is important in the discrimination of, or preference for, faces. Other studies have suggested that movement of the internal features makes faces more discernible and interesting to babies. Vinter (1986) found that infants imitated facial gestures (mouth opening and tongue protrusion) when they were mobile rather than static. However, Legerstee (1991) showed that infants imitated these facial gestures only if they were produced by faces of people, rather than by objects simulating the gestures. This means that not just any movement facilitated discrimination and imitation of facial gestures. Rather, to enhance discrimination and social responsiveness in infants, the movement needs to have a coherent structure. A study by Stucki, Kaufmann-Hayoz, and Kaufmann (1987) supports this interpretation. The authors found that 3-month-old infants were able to differentiate moving white spots painted on a black human face from those painted on a black nonhuman mask only if the face was presented the right way up.

The Stucki et al. (1987) study does not clarify whether infants responded to a familiar kinetic movement pattern or a social stimulus, because the two are confounded in that study. A moving face is both a social and a familiar moving pattern. Recall that the infants looked longer at the doll than at the peer, and least long at their own face, but that they reserved their smiles and vocalizations only for people. This suggests different bases for the responses; looking may be a result of novelty, whereas smiles and vocalizations have a potential social significance. It could be argued, however, that infants smiled and vocalized more often to social faces than to the nonsocial faces because of perceptual discrepancy and not because they recognized the social faces as a communicative or social partner. The faces of the infants were more similar in appearance than either was to the faces of the nonsocial objects. Even though the

This content downloaded from 148.197.98.12 on Wed, 2 Oct 2013 12:19:43 PM
All use subject to JSTOR Terms and Conditions
babies discriminated between the two human faces with significantly different looking patterns, they may have found them more attractive than the dolls' faces and therefore smiled and vocalized more to these images than to the nonsocial images. We would need to explore the role of attractiveness of human faces in affective communication more fully.

Whatever the basis for the early discrimination among self, peer, and object, and between social and nonsocial stimuli, its appearance during the first 5 months of life is likely to have important implications for early social development. Although discrimination may not indicate that infants recognize themselves as social entities, it certainly forms a perceptual basis on which such knowledge is built. For instance, Bahrick et al. (in press) suggested that infants' abilities to distinguish between their own faces and those of peers may be related to their prior visual experience with their mirror images because all infants experienced daily mirror exposure. Mirror exposure allows infants to become aware of their own movements as they are specified both kinesthetically and visually. However, infants do not view their faces in mirrors in isolation, but with their caregivers while playing imitative and other socio/affective games with them (Fogel, 1995; Stern, 1995). As a result of the adults' responses to their infants' socioaffective behaviors, infants learn to identify with people and to differentiate themselves from others (Tomasello, 1995). Thus one may speculate that, when infants have learned to recognize their distinctive facial features as familiar through exposure to mirror images, they simultaneously have learned to recognize their faces as socio/affective stimuli. It would seem that this sense of self is developmentally the first sense of a social or interpersonal self (Neisser, 1995; Tomasello, 1995).

The findings that infants respond with social responses to self and peer, but not to objects, replicates existing data on infants' differential responsiveness to social and nonsocial objects (Ellsworth et al., 1993; Legerstee, 1991, 1994, 1997b; Legerstee et al., 1987, 1990). However, our infants only smiled and vocalized 7%-8% of the time to the faces and 4%-5% of the time to voices. This is less than the 20% of smiles and vocalizations produced when televised or live adults respond to the smiles and vocalizations of infants. The 50% reduction in responsiveness, however, is congruent with other studies using a noncontingent procedure (Legerstee, 1997b; Hains & Muir, 1996). Even though the magnitude of the infant responses was low, our manipulations still produced clear differences.

Based on the work by Martin and Clark (1982), we hypothesized that infants would differentiate between the sounds of self and peer. Interestingly, the infants became more visually attentive when hearing the sound of the peer than when hearing the sound of self. This difference in looking suggests that infants perceived their own vocalizations as familiar and concentrated on identifying the novel sound of the peer. Infants also may have looked longer toward the sound of the peer because they did not recognize it and were waiting for the appearance of the peer's face.

Surprisingly, however, and in contrast to the neonautes in the Martin and Clark (1982) study, our infants vocalized more when hearing their own sounds than when hearing the peers' sounds. There are possible explanations for these results. Infant vocalizations are elicited by stimuli that are interesting to babies (Rosenthal, 1982). During their early solitary vocal play (Weir, 1962) infants may begin to appreciate their own sounds because they enjoy the proprioceptive feeling of making them (Rochat & Morgan, 1995). This pleasurable feeling, paired with the auditory feedback, may stimulate infants to keep vocalizing (S. Trehub, personal communication, February 2, 1996). There is an adaptive reason as well why infants vocalize in response to their own voices. Two- to 6-month-old hearing infants produce vocalizations that match the intonation of what they have just heard (Masataka, 1982), but deaf infants, who are not able to hear human sounds, begin to fall behind hearing infants in their ability to produce speech sounds (Oller & Eilers, 1988). To match the tonal qualities of their language, infants have to actively monitor their own productions and compare them against others (S. Trehub, personal communication, February 2, 1996). This reasoning may also explain why infants in the present study vocalized and smiled more when hearing human sounds than when hearing the nonsocial sounds. They not only discriminated between the vocalizations and nonsocial sounds because of the physical differences, but because they recognized these sounds as familiar and similar to sounds they produce themselves. Thus, the infants' responses were species-, peer-, and self-specific.

There are various reasons why infants have an early awareness of their own faces and voices. From the beginning of life, infants prefer human sounds over nonsocial sounds, but they especially like the high-pitched sounds found in Motherese and in the vocalizations of babies (Werker & McLeod, 1989). Consequently, infants are attracted naturally to the sounds produced by babies. Also from birth, infants are attentive to human faces, but they are captivated by baby faces (Field, 1979; Fogel, 1979). Infants also
perceive information that specifies their relationship with the environment; they can distinguish between stimuli that are “outside” and “inside” the self. Such early sensory experience also contains part of the origins of self (Gibson, 1995, 1979). The daily experience infants have with their dynamic mirror image may result in recognition through the detection of invariant visual-proprioceptive relations that specify the self from mirror stimulation. However, “while many organisms and young infants clearly perceive the contingency between their own behavior and its outcome, it seems that only the higher primates have sufficiently elaborated concepts of causality to attribute their own visual images to themselves” (Butterworth, 1990). Various researchers have formulated theories and provided empirical support for a domain-specific predisposition in infants to perceive causality (Gelman et al., 1995; Leslie, 1984; Michotte, 1963; Premack, 1990; Spelke et al., 1995). Although this domain-specific ability may draw the infant’s attention to the various physical properties of people and self, it is through the dynamics of social interaction with their caregivers that infants become aware not only of their physical or ecological self but also of their social or interpersonal self (Fogel, 1993; Legerstee, in press; Neisser, 1995; Stern, 1995).

ACKNOWLEDGMENTS

This research was funded by a grant from the Social Sciences and Humanities Research Council of Canada (410-94-1743 and 410-96-1743) to M. Legerstee. Portions of this research were presented at the 10th Biennial International Conference on Infant Studies, Rhode Island, Providence, April 1996. We would like to thank Ruth Catney, Leann Morrow, Wendy Perry, Christina Ricciuti, Johanna Robertshaw, and Maria Vitti for helping to conduct the experiments and for coding, Maria Vitti for preparation of the auditory tapes, and the North York General Hospital for its cooperation in recruitment of the participants. Special thanks go to the mothers and babies who so generously donated their time and energy to participate in this study.

REFERENCES


Gallup, G. G. (1982). Self-awareness and the emergence of...


multidisciplinary debate (pp. 44–78) Oxford: Clarendon Press.


