Developmental Psychology
2012, Vol. 48, No. 1, 1–9

0012-1649/11/$12.00 DOI: 10.1037/a0025263

The Role of Intersensory Redundancy in the Emergence of Social Referencing in 5½-Month-Old Infants

Mariana Vaillant-Molina and Lorraine E. Bahrick
Florida International University

Early evidence of social referencing was examined in 5½-month-old infants. Infants were habituated to 2 films of moving toys, one toy eliciting a woman's positive emotional expression and the other eliciting a negative expression under conditions of bimodal (audiovisual) or unimodal visual (silent) speech. It was predicted that intersensory redundancy provided by audiovisual (but not available in unimodal visual) events would enhance detection of the relation between emotional expressions and the corresponding toy. Consistent with predictions, only infants who received bimodal, audiovisual events detected a change in the affect–object relations, showing increased looking during a switch test in which the toy–affect pairing was reversed. Moreover, in a subsequent live preference test, they preferentially touched the 3-dimensional toy previously paired with the positive expression. These findings suggest social referencing emerges by 5½ months in the context of intersensory redundancy provided by dynamic multimodal stimulation and that even 5½-month-old infants demonstrate preferences for 3-dimensional objects on the basis of affective information depicted in videotaped events.

Keywords: infant development, intersensory perception, social referencing, emotion perception

Infant responsiveness to others' emotional expressions is a cornerstone of social and cognitive development. Social referencing, a fundamental social cognitive milestone, refers to the ability to perceive the relation between another's emotional response and the object or event to which it refers, and to subsequently regulate one's exploratory behavior toward the object or event in accordance with the emotional expression (see Campos & Sternberg, 1981; Feinman, 1982; Feinman & Lewis, 1983; Sorce, Emde, Campos, & Klinnert, 1985; Walden & Ogan, 1988). Infants thus learn by observing others' affective responses, moving toward and exploring novel objects that elicit positive emotional expressions and inhibiting exploration of novel objects that elicit negative emotional expressions.

Research demonstrates that infants show reliable evidence of social referencing from about 10–18 months of age. They avoid novel or ambiguous objects and events when adults display a fearful expression toward the objects and approach when adults display a happy expression (Boccia & Campos, 1989; Feinman, 1982; Feinman & Lewis, 1983; Klinnert, 1984; Klinnert, Emde, Butterfield, & Campos, 1986; Mumme & Fernald, 2003; Walden & Ogan, 1988). However, little research has explored the precursors to this ability in early infancy (but see Cleveland, Schug, & Striano, 2007; Striano & Bertin, 2005), and thus it is not known when or under what conditions it emerges.

Social referencing involves a variety of complex social–cognitive skills. It requires infants to selectively attend to and perceive another's emotional expression (affect perception), selectively attend to and perceive the object, person, or event to which attention is directed (object perception), share attention with another person about the object, person, or event (joint visual attention), relate the emotional expression with the object to which it refers (detect affect–object relations), and finally, regulate their behavior toward the object in accordance with the emoter's expression (regulate behavior).

Research indicates that infants show many of these component skills within the first half year of life. Infants can discriminate multimodal (i.e., audiovisual) emotional expressions in familiar adults or contexts by 3–4 months of age (Kahana-Kalman & Walker-Andrews, 2001; Montague & Walker-Andrews, 2001) and in unfamiliar adults by 5–7 months of age (Flom & Bahrick, 2007; Walker-Andrews, 1986). In unfamiliar adults, infants first perceive emotional expressions in dynamic audiovisual speech (3–5 months), later in vocal expressions alone (5 months), and, finally, in facial expressions alone (7 months; Flom & Bahrick, 2007; see Walker-Andrews, 1997, for a review). Infants show gaze following and sensitivity to triadic interactions by 3–6 months of age (Reid,

This article was published Online First September 5, 2011.

Mariana Vaillant-Molina and Lorraine E. Bahrick, Department of Psychology, Florida International University.

This research was supported by National Institute of Mental Health Grant R01 MH 62226, National Institute of Child Health & Human Development Grants R03 HD 052602, R01 HD 053776, and K02 HD06493, and National Science Foundation Grant CSLC SBE 035020 awarded to the second author. The first author was supported by National Institute of Health/National Institute of General Medical Sciences Grant R25 GM061347.

Portions of this data were reported at the 2005 and 2007 biennial meetings of the Society for Research in Child Development held in Atlanta, Georgia, and Boston, Massachusetts, respectively. We thank Laura Batista-Taran, Irina Castellanos, and Melissa Argumosa for their assistance in subject recruitment, testing, and data entry and the infants and parents for their participation.

Correspondence concerning this article should be addressed to Mariana Vaillant-Molina, Department of Psychology, Florida International University, Miami, FL 33199. E-mail: vaillantmolina@gmail.com

Striano, Kauffman, & Johnson, 2004; Striano & Bertin, 2005; Striano & Stahl, 2005). For example, 3- and 6-month-olds have been found to gaze longer and smile more at an experimenter in both dyadic and triadic (experimenter, infant, and toy) contexts when the experimenter coordinated visual attention and affect between an object and the infant than when she did not (Striano & Stahl, 2005). Further, 4- and 6-month-old infants look more toward objects that are the focus of an adult's gaze (Morales, Mundy, & Rojas, 1998; Reid & Striano, 2005). Thus, by the first half year of life, infants detect and discriminate several emotional expressions, follow another's direction of gaze, even within a triadic context, and demonstrate sensitivity to coordinated visual attention and affect during a joint attention episode, all fundamental skills underlying social referencing.

Social referencing has typically been studied in somewhat older infants. Studies have assessed infant regulation of exploratory behavior toward ambiguous or novel objects or events such as toys, a stranger, or a visual cliff, as a function of an adult's positive versus negative emotional expression (Campos & Sternberg, 1981; Feinman & Lewis, 1983; Hertenstein & Campos, 2004; Kim, Walden, & Knieps, 2010; Klinnert et al., 1986; Moses, Baldwin, Rosicky, & Tidball, 2001; Mumme, Fernald, & Herrera, 1996; Sorce et al., 1985). These studies have found reliable evidence of social referencing after infants reach the age of 11–12 months. Only a few studies have tested younger infants. Walden and Ogan (1988) found evidence of behavioral inhibition toward toys following expressions of fear at 10–22 months but not at 6–9 months in live interactions, and Mumme and Fernald (2003) found evidence of behavioral inhibition at 12 but not at 10 months of age following televised audiovisual expressions of fear. The youngest infants to show behavioral regulation thus far appear to be 8½- to 10-month-olds (Boccia & Campos, 1989). They were less friendly toward strangers following expressions of worry than happiness in live interactions. It is not yet known how younger infants would perform in social referencing tasks. Social referencing tasks often require infants to locomote toward (or away from) the referent object, making it difficult to assess prelocomotor infants (younger than 7 or 8 months old). Detection of affect–object relations and perhaps regulation of behavior in accordance with such relations may thus develop earlier than social referencing studies suggest when such measures of exploration as reaching or touching are used, particularly in familiar, dynamic, multimodal contexts.

Most social events occur in dynamic, multimodal contexts (involving visual, auditory, and tactile stimulation) with affective expressions conveyed by synchronous faces and voices. The intersensory redundancy hypothesis (IRH; Bahrick & Lickliter, 2000, 2002), a theory of selective attention, makes clear predictions regarding the contexts that promote development of affect perception and social referencing. Perception of affect is based on amodal information such as rhythm, tempo, duration, and intensity patterns. According to the IRH, affect perception should emerge first in dynamic, multimodal stimulation (e.g., audiovisual speech), where detection of amodal information is promoted (Bahrick & Lickliter, 2002, 2004; Bahrick, Lickliter, & Flom, 2004). Later in development, as infants become more experienced perceivers, perception of affect should extend to unimodal (visual or auditory) stimulation. This developmental progression was supported by recent research findings (Flom & Bahrick, 2007) that demonstrated that by 4 months of age, infants discriminated affective expressions in multimodal (audiovisual) but not unimodal (auditory or visual) speech. However, by 7 months, infants discriminated the same expressions in both unimodal and multimodal speech. Similarly, Walker-Andrews (1997) concluded from a systematic review of the developmental literature that emotion perception emerged first in multimodal stimulation and later extended to unimodal stimulation. Thus, multimodal, dynamic stimulation likely provides the earliest context in which social referencing emerges.

A few studies have examined the role of multisensory information in infants' social referencing. For example, 12-month-old, but not 10-month-old, infants differentially looked at and touched objects that were referenced with a positive expression as opposed to a negative or neutral expression by an actress in a multimodal (audiovisual) televised presentation (Mumme & Fernald, 2003). Similarly, Hertenstein and Campos (2004) found that 11- and 14-month-old infants who viewed live multimodal affective expressions directed toward toys, inhibited play with toys paired with negative compared with positive expressions, even after a delay of 3 min or 1 hr (11- and 14-month-olds, respectively). Vaish and Striano (2004) assessed the role of facial and vocal cues in social referencing in 12-month-olds. While looking at a visual cliff, mothers conveyed happy affect using multimodal (facial and vocal expressions), unimodal visual (facial only), and unimodal auditory (vocal only) expressions. Infants crossed the visual cliff in all conditions but crossed faster and looked longer toward their mothers in multimodal and unimodal auditory (vocal expression) conditions. The authors concluded that by 12 months of age, the voice alone was sufficient for social referencing.

Additionally, Flom and Johnson (in press) assessed 12-month-old infants' detection and memory for affect–object relations when the affective expressions were conveyed by faces and voices (multimodal). Infants were able to detect a change in the affect–object pairing and showed memory for this pairing by preferentially looking at the toy paired with the positive expression up to 1 day later if the experimenter's gaze had been directed toward the object. However, infants showed no evidence of memory for the affect–object pairing if the experimenter's gaze had been directed away from the object. Thus, multimodal presentation of affective expressions, as well as gaze congruence with the direction of the referent object, provides important contexts for infants' perception of affect–object relations. Unlike most previous studies (Vaish & Striano, 2004, is an exception), the present study examined the contribution of multisensory information to infants' detection of affect–object relations and their preferences for objects paired with positive expressions by comparing infants' performance under multimodal and unimodal visual conditions.

Although previous studies provide important information about infants' social referencing skills in unimodal and multimodal contexts, it is not known when social referencing emerges and what conditions promote its development. To this end, the present study assessed the early emergence of social referencing in 5½-month-old infants. Infants' ability to relate two toys with the emotional expressions they elicited in an actress was examined in bimodal (audiovisual speech) and unimodal (visual speech) contexts. Subsequently, infants' touching preferences for the toys was measured. Our present study differed from prior studies because we tested younger infants and combined, in a single design, several conditions thought to maximize the opportunity for social referencing. First, consistent with predictions of the IRH, we expected that redundancy between the face and voice in multimodal speech would facilitate social referencing over unimodal visual stimulation. Thus, to test this hypothesis, we included both unimodal and multimodal conditions. Second, an infant-control habituation procedure was used, providing longer exposure times (averaging between 2 and 4 min for dynamic events; see Gogate & Bahrick, 2001; Bahrick, Hernandez-Reif, & Flom. 2005) to the affect–object pairings than is typical for social referencing tasks in the literature (30 s, Hertenstein & Campos, 2004; and 20 s, Mumme & Fernald, 2003; Vaish & Striano, 2004). Third, the objects used were colorful toys that moved and were designed to engage young infants' attention. And finally, the actress highlighted the affect–object relations by making her gaze and expressions contingent upon movements of the toys, likely maximizing infants' attention to the affect–object relations. Each time the toy moved, she gazed at the toy, responded with an affective expression, then looked at the camera (infant) and back to the toy again.

Contingencies between behaviors and environmental events provide powerful and salient information about the temporal, spatial, and causal relations among events. Perception of contingency is critical for infants' developing communication skills, sense of competence and agency, and ability to predict others' behavior (Bigelow & DeCoste, 2003; Gergely & Watson, 1999; Harrist & Waugh, 2002; Sullivan, Lewis, & Alessandri, 1992; Tarabulsy, Tessier, & Kappas, 1996; Watson, 1972). Contingent events are highly salient to young infants and promote the development of social cognition (Rochat & Striano, 1999). By age 3–5 months, infants can detect the temporal and spatial contingencies provided by video feedback from their own leg and arm movements (Bahrick & Watson, 1985; Rochat & Morgan, 1995; Schmuckler, 1996). Young infants are also sensitive to contingency in face-to-face interactions (Beebe et al., 2010; Bigelow & DeCoste, 2003; Bigelow, MacLean, & McDonald, 1996; Hains & Muir, 1996; Harrist & Waugh, 2002) and prefer contingent over noncontingent responses (Bigelow & Birch, 1999; Murray & Trevarthan, 1985). Young infants can also detect environmental contingencies including forming spatiotemporal expectations by anticipating the timing, location, and appearance of briefly presented images (Haith, Hazan, & Goodman, 1988; Haith & McCarty, 1990; Wentworth & Haith, 1992). In the present study, we created events depicting the natural contingent structure between actions of objects and the contingent gaze and affective expressions they elicited in order to maximize infant attention to the relation between the emotional expressions and the toys.

The present study explored the multimodal origins of social referencing in 5½-month-old infants. Two research questions were posed. First, would bimodal audiovisual stimulation (which provides intersensory redundancy) enhance young infants' ability to relate affective expressions with moving toys to a greater extent than unimodal visual stimulation (which provides no intersensory redundancy)? If so, infants habituated with audiovisual (but not unimodal visual) films of an actress responding with positive affect to the movements of one toy and with negative affect to the movements of another toy should show significant visual recovery to a change in affect–object pairing during the test trials. And second, would infant detection of affect–object relations in videotaped events translate to infant behavioral regulation with respect to the three-dimensional toys? If so, following bimodal audiovisual habituation (but not unimodal visual habituation) infants should show a preference for manipulating the toys that had been previously paired with the positive expression during the habituation phase. This would illustrate that perception of affect-object relations translates to appropriate exploratory behavior toward the object, the essence of social referencing.

Method

Participants

Thirty-two infants (19 boys and 13 girls) with a mean age of 169 days ($SD = 5.2$) participated. All infants had a gestational age of at least 38 weeks and Apgar scores of at least 9 (out of 10). Twenty-eight infants were Hispanic, 3 were White, and 1 was African American. Ten additional infants participated, but their data were excluded due to experimenter error ($N = 1$), or the infants' excessive fussiness ($N = 2$), failure to meet the fatigue criterion (looking at the control stimulus for at least 20% of initial looking time, $N = 3$), or failure to habituate within 20 trials ($N = 4$).

Stimuli and Apparatus

Two toys served as stimulus objects for the study, a robot and a pony. The three-dimensional toys were used for all preference tests, whereas videos of the affect–object events were used for the habituation phase. We created four filmed events depicting each toy moving across a surface eliciting a happy/excited or a fearful/avoidant affective expression from an actress (see Figure 1). Each film consisted of 30-s clips that looped, and each 30-s clip depicted five instances of the affective expressions and five toy movements. The actress's gaze and expressions were contingent on the movements of the toy so that each time the toy moved and made its characteristic sounds, the actress responded by suddenly looking at the toy, turning back to the camera and saying, “Look, it moves. Oh, it moved!” using either a happy/excited or a fearful/avoidant facial and vocal expression. The toy pony moved forward and bobbed its head while neighing. The toy robot moved forward and intermittently opened its arms while making a mechanical clicking noise. Only the surface of the table, the toy, and the head and shoulders of the actress were visible to ensure that the infants' attention to the affect–object relations was maximized (see Figure 1). The pony and robot toys were of similar sizes (7.2 × 6.2 in. [18.29 × 15.75 cm] and 4.8 × 5.8 in. [12.19 × 14.73 cm], respectively) but differed in shape and color/pattern. One version of the filmed events was bimodal (audiovisual), depicting the natural synchrony between sights and sounds. Another version of the events was unimodal (visual) and was created by removing the soundtrack.

During the habituation phase, infants were seated in an infant seat on top of a table facing two 19-in. television monitors, approximately 22 in. [55.88 cm] away. The monitors were surrounded by black poster board with two apertures from which observers could record the infants' visual fixations using joy sticks connected to a computer that recorded and computed the duration of visual fixations to the displays and signaled another experimenter when to commence and terminate each trial. A camera (Digital Handycam Model DCR-TRV30; Sony Corp., Tokyo, Japan) recorded the infants' facial expressions during the habituation

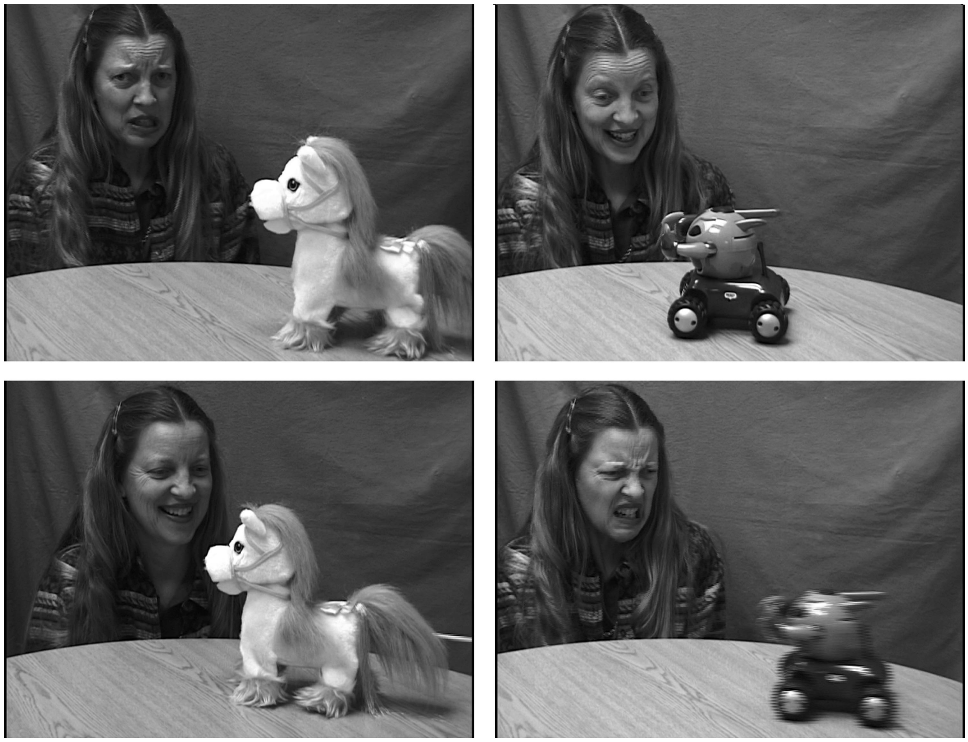

Figure 1. Photos of stimulus events used for the habituation and switch test phases.

phase. During the manual preference test, the three-dimensional toys were placed on a tray in front of the infant, and the camera recorded infants' manual exploration of the toys for subsequent coding.

Procedure

The procedure consisted of an initial visual preference pretest with the three-dimensional toys, followed by an habituation and switch test with the filmed events, followed by a visual preference posttest, and finally a manual preference test with the three-dimensional toys (see Table 1). Infants were randomly assigned to the bimodal or unimodal visual conditions (n = 16 each). The two conditions were identical except that the audiovisual filmed events were used in the habituation and the test phases of the bimodal condition, whereas the unimodal visual (silent) events were used during the habituation and test phases of the unimodal condition. Removing the soundtrack for the unimodal condition eliminated the intersensory redundancy created by the synchronous visual and auditory stimulation.

Visual preference pretest. Parents were asked if their infant had ever seen either of the two toys, and all reported they had not. Experimenters presented the infants with the three-dimensional toys to assess baseline preferences. Infants were seated on a table top and held by an experimenter (who was blind to the hypotheses of the study). A second experimenter, on the right side of the infant, placed the robot and pony toys side by side on a tray

Table 1
Study Procedure

	Phases			
Variable	Visual preference pretest	Habituation & switch test	Visual preference posttest	Manual preference test
Trial structure	Two 10-s trials	Between six and 20 infant-controlled trial	Two 10-s trials	Two 30-s trials
Stimuli	Three-dimensional toys presented side by side	Alternating films of two toys each paired with one affective expression	Three-dimensional toys presented side by side	Three-dimensional toys presented side by side
Conditions	Unimodal visual	Bimodal audiovisual vs. unimodal visual	Unimodal visual	Unimodal visual
Dependent variables	PTLT to each toy	Visual recovery to switch in affect–object pairing	PTLT to each toy	PTNT & PTTT to toy previously paired with happy affect

Note. PTLT = proportion of total looking time; PTNT = proportion of total number of touches; PTTT = proportion of total time spent touching the toy paired with happy affect.

covered with a black cloth. Then, the infant was turned to face the toys, the experimenter uncovered the toys, walked behind the infant, and two 10-s trials began. The toys remained on the tray directly in front of the infant and out of reach during the two test trials. The experimenter covered the toys with the black cloth for a few seconds between trials, while the lateral positions of the toys were switched, and then uncovered the toys, and the second trial commenced. The lateral positions of the toys were also counterbalanced across infants so that half the infants viewed the toy pony on the right and the robot on the left, and half viewed the opposite arrangement first. Infants' looking times to the toys were measured.

Habituation and switch test. Approximately 1 min after the pretest, infants participated in an infant-control habituation procedure where they had the opportunity to learn to relate each emotional expression (e.g., happy/excited vs. negative/fearful) with a particular toy (e.g., robot vs. pony). Learning was assessed in the test trials where the relation between the toy and affective expression was switched. During habituation, half the infants viewed the pony eliciting the actress' happy/excited expression and the robot eliciting the fearful/avoidant expression, whereas the other infants received the opposite pairing. The pony paired with one expression and the robot paired with another expression were shown on alternating trials. Half the infants received the pony first and the robot toy second in the alternating pattern, and half received the opposite order. Infants viewed the alternating films for a minimum of six and a maximum of 20 infant-controlled trials. Each trial was presented until the infant looked away for at least 1.5 s or until the trial reached the maximum of 60 s. A 2-s interstimulus interval followed the end of each trial during which a black screen was displayed. The average of the infants' looking times to the first two trials (baseline) was calculated. Once the infants decreased their looking time to the films by at least 50% of baseline for two consecutive trials, they were considered to have habituated to the events.

After the habituation criterion was met, infants were presented with two additional no-change posthabituation trials followed by two switch test trials. We calculated visual recovery by subtracting looking time to the test trials from looking to the no-change posthabituation trials (see Bahrick, 1992; Bahrick et al., 2005; Flom & Bahrick, 2007). In order to ensure that infants had actually habituated, their looking time to the posthabituation trials was compared with that of the habituation criterion. If posthabituation looking was not less than criterion, infants were rehabituated, receiving additional habituation trials until the habituation criterion was reached again (four of the 32 infants were rehabituated). Then, the two no-change posthabituation trials were repeated. This ensured that all infants had habituated to the two events before being presented with the test trials. Two test trials depicting a switch in the affect–object relations were presented. For example, if an infant had been habituated to the pony eliciting a happy/excited expression and the robot eliciting a fearful/avoidant expression, then the test trials depicted the robot eliciting the happy/excited expression and the pony eliciting the fearful/avoidant expression. It was expected that if infants had detected and learned the relation between the toy and the affective expression, they would show a significant visual recovery to the switch in the affect–object pairing during the test trials. Trained observers unaware of the infant's condition and unable to see the events on the monitor recorded infants' visual fixations. Interobserver reliability was obtained for 25% of the participants.

Visual preference posttest and manual preference test. Following the habituation phase, infants were again presented with the three-dimensional toys. An experimenter placed the toys side by side on a tray approximately 17 cm in front of the infant for two 10-s trials (identical to the visual preference pretest), and infants' looking to both toys was measured. Then, the experimenter covered the toys with a black cloth for a few seconds, switched their lateral positions, pushed the tray toward the infant so that the toys were within reach, and uncovered the toys to begin the first manual preference test trial. Infants were given two 30-s trials to touch and explore the three-dimensional toys. After the first manual preference test trial, the experimenter again covered the toys with the cloth and switched their lateral positions before presenting the second manual preference test trial. Infants' number of touches to the toys as well as the amount of time spent touching each toy was measured from the videotaped recordings and served as the dependent variables.

Results

Visual Preference Pretest

Infants' looking preferences to the toy robot and pony were measured, and interobserver reliability between trained observers was calculated for 25% of the participants using a Pearson product-moment correlation. Interobserver reliability averaged .92 ($SD = .04$). To determine whether infants showed a significant preference for the toys prior to the habituation phase, we conducted a single sample t test on the proportion of total looking time (PTLT) to each toy against the chance value of .5. Results revealed no significant preference ($M = 0.52$, $SD = 0.19$, for robot; $M = 0.48$, $SD = 0.19$, for pony), $t(31) = 0.94$ $p > .1$, $\eta^2 = .03$. Right- and left-looking preferences were also tested and revealed no significant effect ($M = 0.52$, $SD = 0.16$, for right side; $M = 0.48$, $SD = 0.12$, for left side), $t(31) = 0.98$ $p > .1$, $\eta^2 = .03$.

Habituation and Switch Test

Although infants were filmed during the habituation phase for later coding of affective expressions, few displayed detectable affective expressions and thus no analyses were performed. Infants' looking times for each trial type (baseline, posthabituation, test trials, and visual recovery to the change in affect-object pairing) and condition (bimodal vs. unimodal visual) are presented in Table 2. Interobserver reliability between trained observers was calculated for 25% of the participants for looking time to the habituation and switch test events using the Pearson product-moment correlation. Interobserver reliability averaged .99 ($SD = .01$).

In order to address the main research question of under what conditions, if any, did infants detect the switch in affect–object relations, we performed a repeated measures analysis of variance with condition (bimodal audiovisual vs. unimodal visual) as the between-subjects factor and trial type (baseline, posthabituation, and test) as the repeated measure. A significant effect of trial type was found, $F(2, 28) = 19.4$, $p < .01$, $\eta^2 = .58$, and post hoc tests revealed greater looking during baseline than during posthabituation or test ($ps < .01$) and greater looking during test than

Table 2
Means and Standard Deviations for Infants' Looking Times (in Seconds)

	Bimodal audiovisual		Unimodal visual	
Trial type	*M*	*SD*	*M*	*SD*
Baseline	49.4	10.9	29.7	16.5
Posthabituation	7.7	5.7	6.2	3.5
Test	20.7	9.1	7.3	5.5
Visual recovery	13.00*	11.3	1.1	4.6

Note. Baseline is the mean looking time during the first two habituation trials and reflects initial interest level. Posthabituation is the mean looking time during the two no-change trials just after the habituation criterion was met and reflects final interest level for habituation events. Test is the mean looking time during the two switch test trials. Visual recovery is the difference between looking time during the test trials and during the posthabituation trials.
* $p < .01$.

posthabituation trials, $p < .01$. A significant effect of condition was also found, $F(1, 30) = 16.7$, $p < .01$, $\eta^2 = .36$, with greater looking in the bimodal than the unimodal condition. Most relevant to our research question, a significant trial type by condition interaction was found, $F(6, 29) = 13.05$, $p < .01$, $\eta^2 = .73$, with greater looking during the test trials than during posthabituation trials for the bimodal condition, but not for the unimodal condition.

To further explore the condition by trial type interaction (see Table 2) and to determine under which condition infants detected the switch in affect–object pairing, we used single sample *t* tests to evaluate visual recovery scores (increase in looking from posthabituation to test). Infants in the bimodal condition showed significant visual recovery ($M = 13.00$, $SD = 11.26$), $t(15) = 4.8$, $p < .01$, $\eta^2 = .61$, to the change in affect–object relations, whereas those in the unimodal visual condition did not ($M = 1.1$, $SD = 4.64$), $t(15) = 0.88$, $p > .05$, $\eta^2 = .05$. Moreover, visual recovery in the bimodal audiovisual condition was significantly greater than that of the unimodal visual condition, $t(30) = 4.1$, $p < .01$, $\eta^2 = .40$. Thus, infants who were habituated to the bimodal audiovisual events showed evidence of detecting the change in affect–object relations while those habituated to the unimodal visual events did not.

Secondary analyses were also performed to assess any differences in interest in the bimodal versus unimodal events. Two independent sample *t* tests were conducted to determine whether infants in the bimodal versus unimodal conditions differed in their initial interest level (baseline), final interest level (posthabituation), or total looking time (number of seconds looking during habituation) to the habituation stimulus events. Results revealed a significant difference in initial interest level, $t(30) = 3.9$, $p < .05$, $\eta^2 = .34$, with infants in the bimodal audiovisual condition ($M = 49.4$, $SD = 10.9$) looking longer initially to the stimulus events than infants in the unimodal visual condition ($M = 29.7$, $SD = 16.5$). However, there was no significant difference in final interest level ($M = 7.7$, $SD = 5.7$, bimodal condition; and $M = 6.2$, $SD = 3.5$, unimodal condition), $t(30) = 0.857$, $p > .1$, $\eta^2 = .09$, or total looking time to the habituation events across conditions ($M = 302.5$, $SD = 152.1$, bimodal condition; and $M = 222.4$, $SD = 157.9$, unimodal condition), $t(30) = 1.9$, $p > .1$, $\eta^2 = .11$.

Visual Preference Posttest

Analyses were identical to those of the visual preference pretest. Single sample *t* tests revealed no significant PTLTs for the toy previously paired with the happy expression for infants in the bimodal or unimodal condition ($M = 0.55$, $SD = 0.16$, bimodal condition; and $M = 0.56$, $SD = 0.15$, unimodal condition), $ps > .1$. Further, mean PTLTs were compared for the pretest versus posttest, and results revealed no significant shift in preference ($M = 0.03$, $SD = 0.17$, bimodal; and $M = 0.06$, $SD = 0.24$, unimodal), $ps > .1$.

Manual Preference Test

Touches were coded as any manual or oral contact with the toys. On average, infants spent 59% of the total exposure time touching the toys ($M = 0.60$, $SD = 0.29$, bimodal audiovisual condition; $M = 0.58$, $SD = 0.25$, unimodal visual condition) and on average, touches lasted approximately 1.4 s ($M = 1.5$, $SD = 1.7$, bimodal audiovisual condition; and $M = 1.3$, $SD = 1.2$, unimodal visual condition). Touching preferences for the toy placed on the right versus left side were coded. No significant preference for side of presentation was found ($M = 0.51$, $SD = 0.18$ for right side).

The proportion of the total number of touches (PTNT) as well as the proportion of total time spent touching (PTTT) the toy previously paired with the happy expression during the habituation phase served as the primary indices of social referencing. We calculated PTNT for each infant by dividing the infant's total number of touches to the toy previously paired with the happy expression by his or her total number of touches to both toys and averaging across infants. We calculated PTTT for each infant by dividing the total number of seconds he or she spent touching the toy previously paired with the happy expression during the habituation phase by the total number of seconds he or she spent touching both toys and averaging across infants. To address the main research question of whether infants regulate their exploratory behavior toward the three-dimensional objects as a function of the emotional expressions paired with them during habituation, we evaluated PTNT and PTTT with single sample *t* tests. Results demonstrated that the mean PTNT for infants in the bimodal audiovisual condition was significantly greater than the chance value of .5 ($M = 0.62$, $SD = 0.19$), $t(13) = 2.3$, $p < .05$, $\eta^2 = .3$, but the PTNT for infants in the unimodal visual condition was not ($M = 0.53$, $SD = 0.22$), $p > .1$. Further, mean PTTT for infants in the bimodal audiovisual condition was significantly greater than the chance value of .5 ($M = 0.63$, $SD = 0.17$), $t(13) = 2.7$, $p < .05$, $\eta^2 = .3$, but the PTTT for infants in the unimodal visual condition was not ($M = 0.53$, $SD = 0.2$), $p > .1$. Thus, infants who were habituated to the bimodal audiovisual events showed a significant touching preference for the three-dimensional toy previously paired with the happy expression during habituation, whereas infants habituated to the unimodal visual events showed no touching preference for either toy.

Discussion

The present study examined the origins of social referencing in 5½-month-old infants. Infants' ability to relate an actress's positive versus negative emotional response to the movements of two novel toys (a pony and a robot) in a bimodal audiovisual or a unimodal visual video presentation was examined. Following habituation to events that paired the positive expression with one toy and the negative expression with the other toy, infants in the bimodal, audiovisual (but not the unimodal visual) condition showed visual recovery to a switch in the affect–object pairing, indicating they detected the change in the object–affect relations. Then, infants' touching preferences for the toys that had previously been paired with the positive versus negative affective expressions were evaluated to determine whether infants used information acquired during habituation to regulate their exploration of the three-dimensional objects. Results indicated that infants who received bimodal, audiovisual events (but not infants who received unimodal visual events) showed evidence of detecting the affect–object relations and subsequently regulated their manual exploration of the three-dimensional toys, preferentially touching the toy that had previously elicited a happy expression in the actress. These findings indicate that young infants can detect the relation between an affective response to a moving novel toy and use that relation to guide their manual exploration of the toys in multimodal, dynamic contexts. This provides the first demonstration of social referencing in infants as young as 5½ months old.

Results were consistent with predictions generated from the IRH (Bahrick & Lickliter, 2000, 2002; Flom & Bahrick 2007). Intersensory redundancy, available in bimodal audiovisual, but not in unimodal visual, stimulation focuses attention on and promotes perceptual processing of events and their redundantly specified properties, including affective information in audiovisual speech (Flom & Bahrick, 2007). Consequently, detection of affect emerges first in redundant audiovisual speech contexts and is later extended to unimodal auditory and finally to unimodal visual contexts (see Flom & Bahrick, 2007). It was thus predicted that infants receiving bimodal, audiovisual stimulation would show enhanced detection of affect–object relations and would subsequently show significant manual preferences for objects previously paired with the positive affective expression, relative to infants receiving unimodal visual stimulation during habituation. The present findings support our predictions and suggest that social referencing and detection of object–affect relations emerge in rich multimodal contexts where attention and perceptual processing of the relations between affective expressions and the events to which they refer is maximized.

The present findings indicate that by age 5½ months, infants showed a variety of perceptual and cognitive skills fundamental to social referencing. They discriminated happy and fearful affective expressions in naturalistic audiovisual facial and vocal stimulation (affect perception). Infants also discriminated two novel toys, differing in color, shape, and movement (object perception). Moreover, they detected the relation between an affective expression and the object to which it referred in multimodal audiovisual stimulation (detect object–affect relations). They also remembered the affective expressions, the objects, and the relations between them (recall affect–object relations). Finally, infants used their knowledge of the affect–object relations to guide their manual preferences for the toys that had previously been paired with the positive emotional expression in multimodal dynamic displays (regulate behavior).

Although infants did not show preferential looking during the visual preference posttest to the toy previously paired with the positive expression during habituation, they did preferentially touch the toy paired with the positive affective expression. Although infants' looking preferences were in the predicted direction, they were characterized by rather high variability (with *SD*s ranging from 0.15 to 0.24). In contrast with manual preferences, looking preferences may be influenced by a variety of factors and thus provide more indirect measures of preference. For example, one might look at an aversive or potentially dangerous object but not approach or touch it. Accordingly, to evaluate the effects of emotional information from others on infant behavior toward objects, the majority of experimenters in social referencing studies have used more direct measures consisting of infant action toward objects or people.

It is interesting that, consistent with predictions, infants demonstrated touching preferences for the toy previously paired with the happy expression during habituation, despite the fact that infants had viewed a brief reversal of this pairing during the switch test that followed. This apparent lack of interference from the switch test on later preferences is consistent with findings from prior studies in which visual preference tests have followed habituation and switch tests (Bahrick, et, al., 2005; Flom & Johnson, in press; Gogate & Bahrick, 2001). This is likely due to the fact that the habituation paradigm allows infants to fully process both events (e.g., robot paired with the happy expression and pony paired with the fearful expression), whereas the exposure to the reversal during the switch test (e.g., robot paired with the fearful expression and pony paired with the happy expression) is relatively brief and occurs after original learning. Thus, in the present study, infants spent an average of 262.4 s (SD = 172.8) viewing the events during habituation and only 26.51 s (SD = 24.46) during the switch test. Apparently, the more extended processing time for affect–object relations during habituation overrides that of the switch test. Information from the switch test is also likely processed in the context of the prior habituation events, since the switch occurred later and would be experienced as more novel when processed in relation to the habituation events. In any case, manual preferences for the object previously paired with the happy expression in the bimodal condition were robust despite infants having seen a brief reversal of this relation during the switch test.

Why might infants as young as 5½ months demonstrate evidence of social referencing in the present study, when prior studies have typically found evidence after 11–12 months (Boccia & Campos, 1989; Campos & Sternberg, 1981; Feinman & Lewis, 1983; Hertenstein & Campos, 2004; Kim et al., 2010; Klinnert et al., 1986; Moses et al., 1996; Mumme et al., 1996; Sorce et al., 1985), but no evidence in infants younger than 8½ months (Walden & Ogan, 1988). In the present study, we maximized the likelihood of eliciting social referencing through a combination of factors. As discussed earlier, the bimodal audiovisual condition provided intersensory redundancy that maximized attention to the event, the affective information, and the relation between the two. Successful social referencing was observed only when intersensory redundancy was provided by multimodal stimulation at 5½ months of age. In addition, however, there were a number of other variables that were not directly manipulated but that likely contributed to early evidence of social referencing. First, the present study used interesting (colorful, age appropriate) and engaging novel toys. Second, the toys moved in interesting ways, further attracting infants' attention to the events. These factors, along with complete processing afforded by the habituation paradigm, may have made the task easier for young infants. Finally, and perhaps most important, the actress' affective expressions and eye gaze were contingently responsive to the movements of the toys promoting attention to the relation between the affective expressions and the toys. Contingencies are highly salient to young infants (Harrist & Waugh, 2002; Rochat & Striano, 1999), highlighting the relations between the movements of the objects and the contingent affective responses they elicited. Infants are also sensitive to eye gaze direction, and gaze promotes infant attention to objects (Flom & Johnson, in press; Morales et al., 1998; Reid & Striano, 2005). In the present study, the actress' eye gaze shifted to the toy at the onset of its movement and then briefly alternated between the toy and the camera (just after the toy moved) as if looking directly at the participant while vocalizing. For some or all of these reasons, the conditions of the present study elicited evidence of social referencing at age 5½ months, considerably younger than previously observed.

Exploring conditions that elicit evidence of developmental capabilities earlier than previously found can provide an important avenue for uncovering the basis and fundamental conditions that promote the development of the capabilities in question. The present study demonstrates that intersensory redundancy provided by synchronous audiovisual stimulation (across face and voice and across the sights and sounds of object motion) constitute one important condition for supporting the development of social referencing. Further research is needed to determine the relative importance of other factors such as contingency, eye gaze direction, and object motion in the emergence of social referencing.

References

Bahrick, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology, 53,* 180–199. doi:10.1016/0022-0965(92)90048-B

Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face–voice relations. *Developmental Psychology, 41,* 541–552. doi:10.1037/0012-1649.41.3.541

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36,* 190–201. doi:10.1037/0012-1649.36.2.190

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 153–187). New York, NY: Academic Press. doi:10.1016/S0065-2407(02)80041-6

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, and Behavioral Neuroscience, 4,* 137–147. doi:10.3758/CABN.4.2.137

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides infants' selective attention, perceptual, and cognitive development. *Current Directions in Psychological Science, 13,* 99–102. doi:10.1111/j.0963-7214.2004.00283.x

Bahrick, L. E., & Watson, J. S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology, 21,* 963–973. doi:10.1037/0012-1649.21.6.963

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., . . . Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & Human Development, 12,* 3–141. doi:10.1080/14616730903338985

Bigelow, A. E., & Birch, S. A. J. (1999). The effects of contingency on previous interactions on infants' preference for social partners. *Infant Behavior and Development, 22,* 367–382. doi:10.1016/S0163-6383(99)00016-8

Bigelow, A. E., & DeCoste, C. (2003). Sensitivity to social contingency from mothers and strangers in 2-, 4-, and 6-month-old infants. *Infancy, 4,* 111–140. doi:10.1207/S15327078IN0401_6

Bigelow, A. E., MacLean, B. K., & McDonald, D. (1996). Infants' response to live and replay interaction with self and mother. *Merrill-Palmer Quarterly, 42,* 596–611.

Boccia, M., & Campos, J. J. (1989). Maternal emotional signals, social referencing, and infants' reactions to strangers. In N. Eisenberg (Ed.), *Empathy and related emotional responses: New directions for child development* (pp. 25–49). San Francisco, CA: Jossey–Bass.

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

Cleveland, A., Schug, M., & Striano, T. (2007). Joint attention and object learning in 5- and 7-month-old infants. *Infant and Child Development, 16,* 295–306. doi:10.1002/icd.508

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

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

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology, 43,* 238–252. doi:10.1037/0012-1649.43.1.238

Flom, R., & Johnson, S. (in press). The effects of adults' affective expressions and direction of visual gaze on 12-month-olds' visual preferences for an object following a 5-minute, 1-day, or 1-month delay. *British Journal of Developmental Psychology.*

Gergely, G., & Watson, J. S. (1999). Early social–emotional development: Contingency perception and the social–biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first few months of life* (pp. 101–136). Mahwah, NJ: Erlbaum.

Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old infants' memory for arbitrary syllable–object relations. *Infancy, 2,* 219–231. doi:10.1207/S15327078IN0202_7

Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant–adult interactions. *Infant Behavior and Development, 19,* 49–61. doi:10.1016/S0163-6383(96)90043-0

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectations and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59,* 467–479. doi:10.2307/1130325

Haith, M. M., & McCarty, M. E. (1990). Stability of visual expectations at 3 months of age. *Developmental Psychology, 26,* 68–74. doi:10.1037/0012-1649.26.1.68

Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental Review, 22,* 555–592. doi:10.1016/S0273-2297(02)00500-2

Hertenstein, M. J., & Campos, J. J. (2004). The retention effects of an adult's emotional displays on infant behavior. *Child Development, 75,* 595–613. doi:10.1111/j.1467-8624.2004.00695.x

Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Development, 72,* 352–369. doi:10.1111/1467-8624.00283

Kim, G., Walden, T. A., & Knieps, L. J. (2010). Impact and characteristics of positive and fearful emotional messages during infant social referencing. *Infant Behavior and Development, 33,* 189–195. doi:10.1016/j.infbeh.2009.12.009

Klinnert, M. (1984). The regulation of infant behavior by maternal facial expression. *Infant Behavior and Development, 7,* 447–465. doi:10.1016/S0163-6383(84)80005-3

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. doi:10.1037/0012-1649.22.4.427

Montague, D. P. F., & Walker-Andrews, A. S. (2001). Peekaboo: A new look at infants' perception of emotion expressions. *Developmental Psychology, 37,* 826–838. doi:10.1037/0012-1649.37.6.826

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development, 21,* 373–377. doi:10.1016/S0163-6383(98)90014-5

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence of referential understanding in the emotions domain at twelve and eighteen months. *Child Development, 72,* 718–735. doi:10.1111/1467-8624.00311

Mumme, D. L., & Fernald, A. (2003). The infant as onlooker: Learning from emotional reactions observed in a television scenario. *Child Development, 74,* 221–237. doi:10.1111/1467-8624.00532

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development, 67,* 3219–3237. doi:10.2307/1131775

Murray, L., & Trevarthan, C. (1985). Emotional regulation of interaction between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 101–125). Norwood, NJ: Ablex.

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience, 21,* 1763–1766. doi:10.1111/j.1460-9568.2005.03986.x

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *NeuroReport, 15,* 2553–2555.

Rochat, P., & Morgan, R. (1995). Spatial determinants in the perception of self-produced leg movements in 3- to 5-month-old infants. *Developmental Psychology, 31,* 626–636. doi:10.1037/0012-1649.31.4.626

Rochat, P., & Striano, T. (1999). Social cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first few months of life* (pp. 3–34). Mahwah, NJ: Erlbaum.

Schmuckler, M. A. (1996). Visual–proprioceptive intermodal perception in infancy. *Infant Behavior and Development, 19,* 221–232. doi:10.1016/S0163-6383(96)90021-1

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effects on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21,* 195–200. doi:10.1037/0012-1649.21.1.195

Striano, T., & Bertin, E. (2005). Relation among joint attention skills in 5- to 10-month-old infants. *British Journal of Developmental Psychology, 23,* 1–11.

Striano, T., & Stahl, D. (2005). Sensitivity to triadic attention in early infancy. *Developmental Science, 8,* 333–343. doi:10.1111/j.1467-7687.2005.00421.x

Sullivan, M. W., Lewis, M., & Alessandri, S. M. (1992). Cross-age stability in emotional expressions during learning and extinction. *Developmental Psychology, 28,* 58–63. doi:10.1037/0012-1649.28.1.58

Tarabulsy, G. M., Tessier, R., & Kappas, A. (1996). Contingency detection and the contingent organization of behavior in interactions: Implications for socioemotional development in infancy. *Psychological Bulletin, 120,* 25–41. doi:10.1037/0033-2909.120.1.25

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

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development, 59,* 1230–1240. doi:10.2307/1130486

Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology, 22,* 373–377. doi:10.1037/0012-1649.22.3.373

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin, 121,* 437–456. doi:10.1037/0033-2909.121.3.437

Watson, J. S. (1972). Smiling, cooing, and "the game." *Merrill-Palmer Quarterly, 18,* 323–339.

Wentworth, N., & Haith, M. M. (1992). Event-specific expectations of 2- and 3-month-old infants. *Developmental Psychology, 28,* 842–850. doi:10.1037/0012-1649.28.5.842

Received September 30, 2010
Revision received April 11, 2011
Accepted May 11, 2011 ■