

## Infants' Perception of Expressive Behaviors: Differentiation of Multimodal Information

Arlene S. Walker-Andrews

Rutgers, The State University of New Jersey—New Brunswick

The literature on infants' perception of facial and vocal expressions, combined with data from studies on infant-directed speech, mother–infant interaction, and social referencing, supports the view that infants come to recognize the affective expressions of others through a perceptual differentiation process. Recognition of affective expressions changes from a reliance on multimodally presented information to the recognition of vocal expressions and then of facial expressions alone. Face or voice properties become differentiated and discriminated from the whole, standing for the entire emotional expression. Initially, infants detect information that potentially carries the meaning of emotional expressions; only later do infants discriminate and then recognize those expressions. The author reviews data supporting this view and draws parallels between the perceptions of affective expressions and of speech.

An essential task for people living in a social world is interacting with others by reading their emotional expressions and responding adaptively. Early researchers of the development of affect in infancy have emphasized the production of emotional expressions and the socialization of emotions (for a review, see Campos, Barrett, Lamb, Goldsmith, & Stenberg, 1983). More recently, psychologists have focused on the development of perception of emotion, in part because contemporary theories about development of the self, affect, and cognition highlight the significance of interpersonal interactions and emotion perception is important to these interactions (e.g., Hobson, Ouston, & Lee, 1988; Rogoff, 1990; Stern, 1985).

Although the development of emotion perception extends beyond infancy—perhaps throughout the lifespan—this review concentrates on its development during the first year of life. This period is featured because of the dramatic changes in emotion perception competencies that are observed over this period of development (Campos et al., 1983). In addition, the understanding of emotion perception and its early development may be critical to the understanding of development of interpersonal skills during infancy. Furthermore, it may be that infants reared in situations with impoverished affective expression information, such as those, for example, from caregivers with clinical depression (Dawson, 1994; Field, 1994; Gelfand & Tėti, 1990; Tronick & Gianino, 1987), or in contexts where actions and expressions are discrepant may be particularly influenced in their comprehension of expressions (Cummings, 1995). In any case, an understanding of these particular situations would seem to rest on the knowledge of the normative processes of the

development of emotion perception in infancy (Malatesta, Culver, Teasman, & Shepard, 1988).

One can argue that the development of emotion perception depends, in part, on the interplay between maturation and the development of separate sensory systems, as reflected in the psychological processes of detection, discrimination, and recognition (Schiff, 1980; Walker-Andrews, 1988). I argue that, although possessing only rudimentary capacities to detect, discriminate, and recognize others' emotional expressions, the human infant is born well prepared to rapidly develop these competencies during the first year. How psychologists study the development of affect perception depends on how they define the psychological components listed above. *Detection* only indicates that an observer is sensitive and responsive to information (Sekuler & Blake, 1994). That is, the sensory systems are affected by it—it is loud enough to be heard, close or large enough to be seen, and so forth. *Discrimination* refers to the ability to tell the difference between two or more objects or events (Sekuler & Blake, 1994). Thus, for emotional expressions, the observer perceives the difference between a smile, for example, and a frown. As argued by me and by others (e.g., A. J. Caron, 1988; Olson, 1981; Oster, 1981; Walker-Andrews, 1988), *recognition* of an emotional expression implies more than detection and discrimination; it involves the person interpreting how someone else will act based on the "expression" in one's face, voice, and gestures. When one tests adults, one has access to their verbal responses and can infer their recognition of emotional expressions. For infants and nonhuman primates, the assertion that the individual recognizes another's affect must be based on other criteria, including evidence that the expressive behavior reliably evokes specific responses (Campos et al., 1983; Muir, Clifton, & Clarkson, 1989; Spitz & Wolf, 1946). This latter criteria is very stringent, however, because the "specific" infant response that one might expect to follow from specific emotional expressions is not always obvious.

Oster (1981) addressed these concerns by proposing several taxonomic categories for facial expressions. In Oster's terms, infants perceive a facial expression as a "stimulus configura-

---

This article was supported by National Science Foundation Grant SBR-9408993. I thank L. Dickson, G. Gambone, and J. Haviland-Jones for reading drafts of this article and M. Ryder for finding elusive references.

Correspondence concerning this article should be addressed to Arlene S. Walker-Andrews, Department of Psychology, Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08903. Electronic mail may be sent via Internet to arlenewa@rci.rutgers.edu.

tion" if they detect information critical to an adult's judgment of emotion. This information must be independent of irrelevant features or contextual variations, such as visual contrast. Perception of a facial expression as a "sign of emotion" entails responding to another's internal state, indicating that the infant has gathered information about the underlying emotion. The response could take the form of a fixed action pattern (e.g., a startle in response to an angry face), empathic response (e.g., crying in response to another's sad expression), or—in older infants—a more cognitive-based reaction, which encompasses inferences about another's experience (e.g., to offer a toy to comfort a crying child). Finally, infants perceive expressions as "social signals," in Oster's view, when they seem to have expectations about whether particular behaviors will follow or accompany the expression perceived.

In what follows, I articulate a view of the development of emotion perception. This view reflects many of the same concerns and distinctions proposed by Oster (1981) and by ethologists, who study primate behavior. The perspective is that expressions are a form of communication that primates, including humans, have evolved to perceive and use to guide social behavior (see, e.g., Andrew, 1963; Chevalier-Skolnikoff, 1973; Kraut & Johnston, 1979; Sackett, 1966; and van Hooff, 1973). Accordingly, I argue that the major issues are (a) whether and when infants detect the information needed to respond to the primary emotions (Izard, 1979), (b) whether and when they can differentiate among emotional expressions, and (c) whether and when they perceive these expressions as meaningful indicators of another's future behavior and as guides for their own actions (Fridlund, 1991). Like many others (cf. McArthur & Baron, 1983; Meltzoff & Moore, 1993; Muir, Humphrey, & Humphrey, 1994), I argue that emotion perception abilities rest on the same abilities as other types of perception. *Perception* is an active process to obtain information about the world (Dewey, 1896; E. J. Gibson, 1988) that allows one to act efficiently and adaptively (J. J. Gibson, 1979/1986). In formulating and articulating this perspective, I draw on what Johnson and Morton (1991) have called the "prevailing view of the development of face perception" (p. 29), the theory set forth by E. J. Gibson (1969).

### Perception of Emotion as a Special Process

The perspective that emotion perception rests on general perceptual processes is in sharp contrast with current "modularity" views of perception and cognition. In general, modularity approaches distinguish between perception and cognition and promote nativistic views of both. In brief, the brain is said to have a modular organization (Fodor, 1983), meaning that human competencies result from the overlapping activity of many partially independent systems or modules (Neisser, 1994). Several modularity arguments relevant to emotion perception have been advanced. These views make a distinction between the perception of persons and of objects: Separate mechanisms or structures are held responsible for each perception. I outline three theories because they speak most directly to the development of emotion perception.

Neisser (1994) has argued that one modular unit is "interpersonal perception-reactivity," which underlies immediate social interactions with others. It is a separate system from "direct

perception-action," which enables the observer to perceive and act in the local environment. Neisser presumed that recognition is not part of direct perception. Instead, it denotes naming or labeling an object, person, or event. Furthermore, he proposed that object perception and person perception are distinct systems because affect is involved only for the latter.

In a similar vein, Leslie (1994a, 1994b) has submitted that there is a specialized mechanism called a theory of mind module (ToMM), which produces learning about others' mental states. He suggested that information about another's behavior arrives through different sensory channels. To make sense of these data, the ToMM operates postperceptually as an interpretative device, calculating the belief state of the other person. The understanding of another's beliefs and attitudes is separate from other sorts of perception and cognition and relies on domain-specific sets of processes. A natural extension of this view asserts that emotion perception also requires a domain-specific mechanism, perhaps the same ToMM or a theory of emotion module (ToEM; Hobson, 1989; Hobson et al., 1988). Evidence for ToMM usually places its onset at about 3 years of age (Leslie, 1987, 1994a).

Finally, Baron-Cohen (1994) has proposed a detailed model for perception of persons. He asserted that humans have evolved to attribute mental states to others to interpret their behavior and predict action (see also Premack & Woodruff, 1978). Baron-Cohen proposed a dedicated neurocognitive system with four modular components: an intentionality detector, eye direction detector (EDD), shared action mechanism (SAM), and ToMM (Leslie, 1991). Each module comes "on-line" at a different time. The intentionality detector is a primitive, amodal perceptual mechanism that "reads stimuli as volitional" (Baron-Cohen, 1994, p. 516). That is, by 6 months, infants interpret the actions of animate objects as intentional. The EDD switches on somewhat earlier (4 months) and has two functions: to detect the presence of eyes and, in higher primates, to represent eye behavior. SAM begins to function at 9 months when infants begin to participate in a "range of joint visual attention behaviors" (p. 533), whereas ToMM "comes on line in the middle of the second year" (p. 537).

These three theories offer innovative ways to view emotion perception, although they are not without their critics. For example, Muir, Hains, and Symons (1994) have reviewed Baron-Cohen's (1994) theory in detail and suggested that infants "between 3 and 6 months of age do not read mental states into non-human stimuli" (Muir, Hains, et al., 1994, p. 672), as proposed. They found that the contingency of responding (Watson & Ramey, 1972), not visual information, is the compelling cue for intentionality. Furthermore, in regards to EDD, gaze direction may signal the possibility of an interaction but does not serve to maintain it. The bulk of data reviewed by Muir, Hains, et al. suggests that the infant's purported "theory of mind" is more complex than the operation of a set of independent modules. With respect to emotion perception itself, a modular organization may be consistent with domain specificity but not with findings that suggest context effects and interactions between systems, especially across modalities (Thelen & Smith, 1994). Whether modularity theories serve as feasible explanations for development in emotion perception is revisited later.

### Perception of Expressions: A Proposed Sequence

In this article, I outline a different proposal about development of perception of emotional expressions. In brief, I propose

that infants' emotion perception develops in the same manner as their perception of other objects and events. Throughout, I indicate ages when I expect that certain competencies are expressed. The age ranges listed, however, should only be viewed as trends because there are substantial variations in the data, owing to individual differences, methodological issues, stimulus properties, and so forth. At the heart of my proposal is the idea that infants first come to differentiate and recognize social signals by abstracting the meaning that is invariant across multimodal presentations. This view is based on a rich literature that includes findings on many aspects of perceptual development. For example, infants show an early sensitivity to motion (Bertenthal, Proffitt, Spetner, & Thomas, 1985; F. Kaufmann, Stucki, & Kaufmann-Hayoz, 1985; R. Kaufmann & Kaufmann, 1980), visual and acoustic information that specifies the affordances of others (Spelke, 1976; Spelke & Cortelou, 1981), and reciprocal patterns in mother-infant interactions (Stern, 1974; Trevarthen, 1979). By the end of the first year, they also show social referencing, a more complex skill in which others' expressions may be used as information about external events (Klinnert, Campos, Sorce, Emde, & Svejda, 1983).

With respect to relevant experimental evidence, methods selected by researchers are varied because they have tried to discover what infants understand about emotional expressions by using combinations of "converging operations" (Garner, 1981; Walker, 1981). Most of this work focuses on the discrimination of facial expressions. Researchers have also examined infants' face and voice perceptions, their responses to infant-directed (ID) speech, mother-infant interactions, and social referencing.

#### Sensitivity to Information for Affect (Detection)

At birth, the perceptual systems are sensitive to information from the environment, working together as an overall system of preadapted coordination (E. J. Gibson, 1983). For example, infants look longer at faces accompanied by a voice (e.g., Haith, Bergman, & Moore, 1977), and they turn toward soft sounds (e.g., Butterworth & Castillo, 1976; Wertheimer, 1961). In addition, infants preferentially attend to persons and other animated objects (e.g., Bertenthal et al., 1985; Johnson, Dziurawiec, Bartrip, & Morton, 1992; Johnson, Dziurawiec, Ellis, & Morton, 1991; Sherrod, 1979). However, sensory systems are restricted as well. For example, both visual acuity and scanning are limited at birth. The fovea holds visual receptors that are neither as developed nor as densely packed as in the adult. By about 3 months, however, the infant's ability to focus accurately approaches an adult's performance (Banks, 1980). By 6 months, visual acuity has improved substantially (Gwiazda, Bauer, & Held, 1989), and contrast sensitivity is sufficient to detect most static facial expression contrasts (Hainline & Abramov, 1992).

Thus, in the first few months, infants are apparently sensitive to perceptual information, which potentially specifies particular emotions, but they respond in modest ways. For example, an infant may gaze at a toothy smiling face, but his or her preference may originate in a tendency to fixate on high-contrast patterns or reflect that particular infant's customary experience with smiles. Proffitt and Bertenthal (1990) pointed out that "demonstrating a common sensitivity to stimulus information does not necessarily imply that adults and infants share meanings" (p. 2).

#### *Perception of Human Voices*

Many studies attest to infants' sensitivity to auditory information, such as frequency, intensity, and temporal structure. In addition, infants seem to treat human voices as distinct (Butterfield & Siperstein, 1970; Ecklund-Flores & Turkewitz, 1996; Hutt, Hutt, Leonard, von Bermuth, & Muntjewerff, 1968). Indeed, one of the most exciting discoveries about neonates' (birth to 1 month old) perception was reported by DeCasper and Fifer (1980) more than 15 years ago. Neonates, only 3 days old with fewer than 12 hr of maternal contact, preferred their mother's voice over the voice of another neonate's mother. DeCasper and Spence (1986) asked women to recite a passage in their last 6 weeks of pregnancy to demonstrate that the reinforcing value of the mother's voice, and even a particular passage, was acquired through prenatal experience. In subsequent studies, researchers have upheld and extended these results (e.g., Ecklund-Flores & Turkewitz, 1996; Moon, Cooper, & Fifer, 1993). Infants may prefer particular voices because of early experience, intrinsically attractive features, or both (Cooper & Aslin, 1990; Fernald, 1985; Turkewitz, Birch, & Cooper, 1972).

#### *Perception of Human Faces*

Neonates' ability to discern fine detail is limited, but within the first few months, they can detect differences in the shapes of objects or two-dimensional figures, perceive spatial relationships, and treat an object viewed from different angles or distances as the same (Bower, 1966; Slater, Mattock, & Brown, 1990). Neonates inevitably look preferentially at patterned visual stimuli, high-contrast patterns, and moving figures. Face perception appears to be most advanced. This may be because visually preferred attributes typify the human face. Alternatively, face perception may be mediated by special purpose mechanisms, although this hypothesis has not been confirmed (Bruyer et al., 1983; Habib, 1986; Morton & Johnson, 1991; Sergent, 1987). In any case, within hours after birth, neonates look longer to facelike stimuli (Goren, Sarty, & Wu, 1975; Johnson et al., 1991; Maurer & Young, 1983). Neonates visually track such materials farther, compared with drawings that contain many of the same features. Later, infants seem to require more realistic representations for the same preference. In part, this may be because of improved acuity. As details become clearer, infants detect more differentiating features, and poorer facsimiles are not adequate (Johnson et al., 1992).

Not only do neonates look preferentially to faces but also gaze longer at their mother's face after brief exposures to it. Four-day-old neonates look longer at their mother's face (Bushnell, Sai, & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984), even in a videotape (Walton, Bower, & Bower, 1992). This is not true when the mother's hairline is obscured, although 4-month-olds are unaffected by such alterations (Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995). Pascalis et al. argued that different perceptual structures are responsible. Whatever the root cause, it seems that neonates often attend to specific features rather than configurational information.

To summarize, infants detect information that may allow for the eventual recognition and discrimination of emotion. What is detected changes rapidly, as perceptual and motor systems develop. With respect to vision, a neonate can just discern a

blurry face and distinguish the hairline, eyes, nose, and mouth (Banks & Ginsberg, 1985). Therefore, it is unlikely that neonates discriminate static facial expressions based on anything other than feature information. Dynamic expressions may provide additional information because infants are sensitive to motion (Fox & McDaniel, 1982), motion has influence on the externality effect (Bushnell, 1979; Milewski, 1976), and affective information is present in the transitions available in dynamic portrayals (Ekman & Friesen, 1978). Moreover, within just a few months, infants can detect additional details, such as wrinkles, laugh lines, and other feature and relational information that mark particular facial expressions. Such perceptual development coincides with improvements in other sensory systems and in accord with cognitive advances (Turkewitz & Devenny, 1993).

## Discrimination of Facial and Vocal Expressions

### *Visual Preference for Facial Expressions*

Early attempts to determine whether infants can discriminate facial expressions typically involved the presentation of static faces in different affective poses, either in a paired preference or successive presentation procedure (e.g., Barrera & Maurer, 1981; Field, Woodson, Greenberg, & Cohen, 1982; LaBarbera, Izard, Vietze, & Parisi, 1976; Young-Browne, Rosenfeld, & Horowitz, 1977). For example, Wilcox and Clayton (1968) gave successive presentations—28- or 60-s films of nodding or static smiling, neutral, and frowning faces—to 5-month-olds in two separate experiments. For the 28-s films, infants gazed longest at the smiling faces. In the second experiment, infants looked at each moving face about 80% of the time available, but no preferences emerged for any facial expression. LaBarbera et al. found that 4-month-olds discriminated between photographed joy and anger and between joy and a neutral face, looking longer to the joy expression. Kuchuk, Vibbert, and Bornstein (1986) asked whether infants could discriminate various intensities of an expression from another facial expression. Three-month-olds showed a preference for a smiling face over a neutral one, especially as the smile became more pronounced, suggesting that the mouth itself was the distinguishing feature.

### *Operant Preference Techniques*

Everhart and Henry (1992) have reported discrimination of facial expressions by neonates 23 to 93 hr old using an operant procedure. Neonates could, by sucking on a pacifier, replace a digitized image of one expression with another (neutral to happy, sad, or angry). Neonates sucked more often to generate the happy expression. It is unclear what information the neonates were using, however.

### *Habituation to Static Facial Expressions*

Another commonly used method is the visual habituation technique. Although the purpose was to examine something more than discrimination, "the infant's response provide[d] direct evidence for nothing more than a discrimination" (Proffitt & Bertenthal, 1990, p. 2). In one study, Barrera and Maurer (1981) reported that 3-month-olds discriminated between frowning and smiling, especially as posed by their own mother.

Young-Browne et al. (1977) found that 3-month-olds could discriminate between photographed happy and surprised expressions. Results were less clear regarding surprised and sad expressions. After habituating to sadness, infants increased fixation to surprise. Yet they did not increase fixation to sad after habituating to surprise. Such order effects are not uncommon, but difficult to explain, as discussed below.

Nelson, Morse, and Leavitt (1979) made an ingenious use of habituation, asking whether infants could categorize facial expressions. Seven-month-old infants were visually habituated to expressions, but several different persons posed each expression during familiarization to introduce variability. Infants discriminated happy and fearful expressions but only if happy expressions occurred first. Nelson and Dolgin (1985) replicated this result using habituation in one part of their study and visual preference in another to isolate the source of the order effect. Infants looked much longer to fear on the preference portion.

Additional evidence for generalization and discrimination of facial expressions has been reported by Serrano and his colleagues. In one study, 4- and 6-month-olds were habituated to anger, fear, and surprise (Serrano, Iglesias, & Loeches, 1992). Each infant viewed three different models, expressing the same emotion on the habituation trials. On these trials, fear was looked at less than was either anger or surprise. On the test trials, infants increased fixation to the novel expression, except when fear followed surprise.

Overall, these data suggest that infants as young as 3–4 months old can discriminate between photographs of different facial expressions, notwithstanding the results of Everhart and Henry (1992). What is not apparent is what kind of information the infants used and what the facial expressions meant to them (cf. Nelson, 1987). Are infants using the affective meaning (e.g., happiness vs. sadness) to discriminate (recognition, in my terminology), or are they using features such as degree of eye widening? Moreover, are the order effects simply artifactual or do they reveal something significant about the infants' abilities? Finally, because expression preferences were inconsistent (e.g., LaBarbera et al., 1976; Nelson & Dolgin, 1985; Serrano et al., 1992), fixation duration cannot be presumed a sensitive measure of partiality.

R. F. Caron, Caron, and Myers (1985) used specially photographed facial expressions with infants ranging from 4 to 7 months to tackle these questions. Eight different women posed "toothy angry," "nontoothy angry," or "nontoothy smiling" expressions. On the posttests, infants saw two novel women posing the habituated expression and then the same two women posing a novel expression, "toothy smiling." R. F. Caron et al. reasoned that, if infants were using affective meaning for discrimination, then only infants in the angry conditions would increase fixation to the apparently novel toothy smiling. Instead, infants at all ages dishabituated if they had habituated to one of the nontoothy expressions, regardless of whether it depicted anger or happiness. Analogously, Oster and Ewy (1980, cited in Oster, 1981) reported that infants looked preferentially to toothy but not closed-mouth smiles when compared with sad faces. In the Caron et al. study, when feature information was held constant across expressions, even 7-month-olds did not discriminate between the expressions.

More recently, researchers have asked whether infants can be induced to focus on the affective information, ignoring feature

information. The feature information presented for habituation differed systematically. In some photographs, people had toothy mouths; in others they did not. Some were squinting, others had rounded eyes, and so forth. Usually several different models posed the expressions. The one variable held constant was affective meaning. After habituation, some infants were presented with a new model, who posed a new version of the familiarized expression. If they noticed the common affective meaning, infants should have generalized to the new display of the familiar expression. The remaining infants were presented with a novel expression.

For example, Ludemann and Nelson (1988) responded to questions posed by R. F. Caron et al. (1985) by asking whether 7-month-olds could generalize across intensities of happy, fear, and surprise and discriminate between expressions. The data obtained were marked with several order effects: Infants generalized across (a) happy expressions and discriminated them from fear only when happy expressions were presented first, (b) surprise expressions and discriminated them from fear only when surprise was presented first, and (c) happy expressions and discriminated them from surprise only when happy expressions were presented first; they (d) tended to look longer to fearful expressions. Ludeman and Nelson concluded that infants, influenced by the relative familiarity of expressions (hence the order effects), may categorize a surprise expression as an instance of a happy expression. Infants' difficulty in encoding fear expressions may be because either fear is an aversive expression or they have little familiarity with it.

Later, Kestenbaum and Nelson (1990) tested 7-month-olds' reliance on affectively relevant versus feature-specific information. Infants saw either single or multiple photographed exemplars of a facial expression in either an upright or inverted orientation. They argued that affectively relevant information is orientation specific. For adults, face recognition is impaired when faces are presented upside down (Köhler, 1940; Yin, 1969). In the first experiment, infants generalized across different models' happy expressions and discriminated them from fear and anger but only for upright faces. By themselves, these results suggest that 7-month-olds can use affective meaning to recognize facial expressions. In a second experiment, 7-month-olds were habituated to a single model posing a happy expression. On the posttests, the model's face expressed anger or fear. Infants in both conditions dishabituated to the novel expressions. That inversion did not interfere suggests that infants used feature rather than affective information. Finally, infants were visually habituated to three models with toothy smiles and then were shown a fourth model who posed nontoothy happy and nontoothy angry. Infants in both the upright and inverted conditions dishabituated to both nontoothy expressions, again evidence for feature-based discrimination.

Serrano, Iglesias, and Loeches (1995), following the lead of Ludemann and Nelson (1988), habituated 4- to 6- and 7- to 9-month-olds to moderate-intensity happy, angry, and neutral facial expressions to reduce reliance on the "marked contrast between lips and teeth in the intense versions" (p. 477), although upturned lips characterized the happy expressions and furrowed brows, the angry expressions. For all pairs, infants dishabituated to the novel expression, except for the younger infants when angry and neutral expressions were compared. Serrano et al. also videotaped infants' behavioral responses to

the expressions. Analyses were carried out by subgroup because one third of the infants saw only happy and angry, one third happy and neutral, and one third angry and neutral expression contrasts. In general, positive reactions tended to be more frequent during habituation to happy than to angry expressions, and negative reactions occurred more often to angry than the other two expressions. However, infants habituated to happy and neutral expressions did not show behavioral differences, and those habituated to angry and happy expressions did not differ in the frequency of positive and negative responses to angry expressions. Ludemann (1991) provided a more exacting test by asking whether infants can treat positive facial expressions (surprise and happy) as similar, while discriminating them from negative expressions (e.g., fear). Ten-month-olds, but not 7-month-olds, demonstrated this ability in her experiment.

These results suggest that 7-month-olds base discrimination of static faces on feature information when it is available rather than on affective meaning. When feature information is not salient, older infants are able to discover the common affect among facial expressions and discriminate them from a novel expression (Kestenbaum & Nelson, 1990; Ludemann, 1991; Ludemann & Nelson, 1988). Serrano et al. (1995) reported similar abilities among infants as young as 4 months old for all tested expressions, save angry and neutral. Generalization of emotional expressions heralds the beginning of an ability to recognize photographed emotional expressions.

#### *Preference for Vocal Expressions*

Studies on infants' perception of vocal expressions are relatively scarce. More recent data have come indirectly from studies of infants' perception of ID and adult-directed (AD) speech, to be discussed in a later section. That aside, in a few experiments, researchers have looked at infants' preferences for vocal expressions. For example, Aldridge (1994) tested neonates using an operant-choice sucking procedure and found that they preferred to listen to happy compared with angry and sad voices. That neonates "worked harder" for the happy voice and "avoided" listening to the sad and angry expressions signifies a preference based on emotional content. This is an exciting possibility but speculative because neonates heard only a single syllable (*hi*) throughout. Spectograms, included in Aldridge's study, showed clearly that each token was different in amplitude and frequency. For now, these results are subject to the same analysis as the data on discrimination of toothy smiles from closed-mouth angry expressions.

#### *Visual Habituation to Vocal Expressions*

In a series of experiments, my colleagues and I (Walker-Andrews & Grolnick, 1983; Walker-Andrews & Lennon, 1991) have investigated young infants' discrimination of vocal expressions. Infants were habituated to a visual stimulus (a slide), accompanied by an ongoing recording of a vocal expression (Horowitz, 1974). After habituation, the vocal expression was changed; the slide remained the same. An increase in fixation to the familiar visual stimulus was interpreted as evidence that the infant had noticed a vocal change. Changes from one vocal expression to another occurred at different points in the recording because an infant control procedure was used.

Using this technique, Walker-Andrews and Grolnick (1983) habituated 3- and 5-month-olds to either a woman's sad or happy voice, along with a slide of her face expressing the same affect. After the infants visually habituated, only the voice was changed either to happy or sad, depending on the habituated vocal expression. Three-month-olds who heard a change in vocal expression from sad to happy increased fixation tenfold; those who heard happy to sad showed a nonsignificant increase. The 5-month-olds dishabituated for both orders. They concluded that infants as young as 5 months old, and possibly as young as 3 months old, can discriminate between sad and happy vocal expressions.

Initially, Walker-Andrews and Lennon (1991) thought that any visual stimulus could accompany the vocal expression, but they investigated its role further. For some infants, the visual stimulus was a facial expression slide that affectively matched the voice heard during habituation, the voice heard during the test, or a third vocal expression. A fourth group saw a checkerboard instead. These 5-month-olds discriminated between happy, angry, and sad vocal expressions when a face slide was presented, regardless of whether it affectively matched the habituated voice, the novel voice, or neither. When a checkerboard was shown, infants did not increase fixation. This last finding is surprising, given that much younger infants can distinguish voices differing in pitch and other acoustic features (Culp & Boyd, 1974). Walker-Andrews and Lennon (1991) proposed that the presence of the face acts "as a setting for attending to the affective quality of the voice" (p. 140); infants in the checkerboard condition were not provided such a setting. Younger infants who have discriminated voices while looking at nonface stimuli may have used physical, acoustic properties of the voices rather than affective meaning.

In summary, infants may discriminate vocal and static facial expressions based on feature differences initially and then based on affectively relevant information. Infants as young as 3 months old discriminated some photographs of facial expressions in several studies (e.g., Barrera & Maurer, 1981; Young-Browne et al., 1977). Somewhat later, infants begin to use affective information and can generalize across varied portrayals of the same facial expression (7 months old, Kestenbaum & Nelson, 1990; 4 months old, Serrano et al., 1995). Vocal expressions may be discriminated as expressions of emotion by infants about 5 months old, but only if a facial expression, even if it is static, is also present. Context effects are common, as illustrated by both this latter finding and the order effects mentioned earlier.

## Recognition of Facial and Vocal Expressions

### *Visual Habituation: Dynamic Stimuli*

Despite the methodological advances described above, the use of photographs to test perception of facial expressions is not without its critics. Some have argued that the crucial invariants that carry affective information in naturally occurring, dynamic, and multimodal events are lost in static photographs (R. F. Caron et al., 1985; Eibl-Eibesfeldt, 1970; Ekman & Friesen, 1978; Fogel, 1993). The challenge is to preserve the essential information in the displays. Therefore, several different methods have been devised to investigate infants' recognition of meaning in affective expressions, often with dynamic displays. A. J. Caron, Caron, and MacLean (1988; see also Walker-Andrews,

1985), for example, used the standard habituation procedure but with dynamic expressions that offer constantly changing information. Infants viewed several films of actresses facially and vocally depicting happy, sad, or angry expressions. Four-month-olds dishabituated to the novel, bimodal, dynamically presented expressions for happy and sad contrasts in the happy to sad direction. Only 7-month-olds responded to differences between happy and angry expressions.

Nelson and Horowitz (1983) introduced holograms as displays. Two-month-olds discriminated happy (winking and smiling face) and neutral facial expressions (the blowing of a kiss). According to Nelson and Horowitz, these infants, lacking stereopsis, could use only the motion cues provided. Five-month-olds did not discriminate the holograms from one another, although they did discriminate a stationary hologram from a moving hologram.

Field et al. (1983) described neonates' discrimination of facial expressions, as depicted by a live actress, for habituation. They reported that neonates imitated some facial expressions. These results are interesting in part because of a confound. On the one hand, because the actress was simultaneously holding the neonate, other subtle cues could have influenced the neonates' fixation. On the other hand, this provided proprioceptive and tactual information to the neonates.

### *Intermodal Preference Method*

The intermodal preference technique (Spelke, 1976) is also been used to ask whether infants recognize emotional expressions. In these experiments, infants must detect the correspondence between information presented to vision and audition. Typically, infants concurrently view two visual displays, accompanied by a single soundtrack corresponding to one of them. By altering the displays in specific ways, the experimenter can limit the information common to the two modalities.

In the first series of experiments using this technique, I (Walker, 1982; Walker-Andrews, 1986, 1988) tested infants, ranging in age from 2 to 7 months. Infants were presented simultaneously with two filmed dynamic facial expressions (from the set happy, sad, neutral, and angry), accompanied by a single vocal expression that affectively matched one facial expression. Two-month-olds looked almost exclusively at the happy expressions, regardless of which vocal expression was played. Four-month-olds increased their looking time to the happy facial expression when it was sound matched, and 5- and 7-month-olds increased fixation to any of the facial expressions that were sound matched. Seven-month-olds who saw happy and neutral films that were presented silently did not show looking preferences.

In one experiment, synchrony relations were disrupted by a delay of the soundtrack for 5 s. Dodd (1979) had reported that infants are sensitive to asynchrony (400 ms) between lip movements and speech, so this manipulation was devised to ensure that infants would not make an intermodal match based on synchrony. Infants did not show any visual preference during the first trial; but, on the second trial, the preference for the sound-matched facial expression was strong, increasing steadily over the trial. Apparently, infants detected asynchrony, but this did not prevent the infants from making an intermodal match based on other information.

In another experiment (Walker, 1982), infants viewed angry and happy facial expression films either in an upright or inverted orientation, again accompanied by a single vocal expression (cf. Kestenbaum & Nelson, 1990). Only those who viewed the upright, dynamic facial expressions looked at the sound-matched expression. Temporal synchrony and rate information, still available in the inverted condition, did not influence their preference. In another experiment (Walker-Andrews, 1986), the lower part of the face was occluded (blocked off) so that synchrony between lip movements and vocalizations was not visible. Again, 7-month-olds preferentially looked to the sound-specified film, although 5-month-old infants did not. This suggests that, during the first year of life, infants develop the ability to detect common affect across bimodal, dynamic presentations of an emotional expression. Temporal synchrony between face and voice is neither imperative nor sufficient for matching to occur. The mouth and lower part of the face need not be visible, but an upright face is required.

Soken and Pick (1992) focused on what information infants might be using. Seven-month-olds were presented with videotaped angry and happy facial expressions, accompanied by a single soundtrack that affectively matched one of them. The vocal expressions were recorded separately by a different woman, who repeated a different text. In this way, synchrony and rate information were modified but affect-specific information was retained. In one condition, infants viewed fully illuminated faces. These infants fixated the visual displays equally. In the second condition, infants were presented with "point-light" visual displays (Bassili, 1979). These displays reveal only facial motion (through the movement of luminous spots) and eliminate facial feature information. These infants looked longer at the sound-specified facial expression. In the second experiment, Soken and Pick devised similar displays, but a single woman was videotaped for both. Face-voice synchrony was eliminated through editing, whereas both affect- and event-specific information were preserved. This time, infants in both conditions looked longer to the sound-specified videotapes.

Soken and Pick (1992) concluded that the 7-month-olds were able to detect correspondences between facial and vocal expressions based on affective meaning (even when produced by different people) but that the asynchrony conspicuous in the fully illuminated condition may have interfered with the infants' matching performance. They proposed that 7-month-olds can recognize happy and angry expressions based solely on motion information.

Subsequently, Soken, Pick, Bigbee, Melendez, and Hansen (1992) tested discrimination of positive and negative expressions using intermodal preference. By and large, 7-month-olds looked longer to affectively consonant (although asynchronous) expressions, even when the contrasts had the same valence (happy vs. interest, angry vs. sad). Infants also looked preferentially to happy and interest facial expressions paired with a sad facial expression, irrespective of the vocal expression.

Intermodal preference experiments suggest that, by 7 months old, infants can detect correspondences between facial and vocal displays of affect. Several explanations have been offered for these results—arousal matching, associative learning, matching based on physical features, and extracting a common meaning. First, infants are said to be aroused to some degree by the vocal expression and then look for the similarly arousing visual

stimulus (the affectively matching facial expression). This is a plausible account, although it does not explain some findings: (a) When synchrony was disrupted, infants did not look at the corresponding facial expression during the first 2-min trial (Walker, 1982); (b) infants did not show preferential looking when the accompanying facial expressions were inverted (Walker, 1982); and (c) infants looked preferentially to the sound-matched facial expression when facial expressions with the same valence were used (Soken et al., 1992). A second explanation—infants have already experienced the faces and voices together, thus having learned to associate them automatically—is weakened by the lag in preferential looking to the sound-matched facial expression when synchrony was disrupted (Walker, 1982). Matching based on common features—rhythm, synchrony, rate, and intensity—may be the explanation, although 7-month-olds continued to show intermodal matching (a) when the soundtrack was delayed (Walker, 1982), (b) when only motion information was available (Soken & Pick, 1992), and (c) when the mouth area was hidden (Walker-Andrews, 1986). They did not, however, show matching when the faces were inverted (Walker, 1982). Younger infants (5 months old) have also displayed some intermodal matching but have not been tested in each of the conditions (speech delay and inversion) used with older infants.

### Social Referencing

Adults not only recognize the emotional expressions of others and use that information effectively in ongoing interactions but also use others' expressions as information about external events. Development of this skill, called *social referencing*, requires one to not only detect and discriminate others' expressions but also connect those expressions to environmental events. Given its cognitive requirements, social referencing would not be expected to emerge until late in the first year (Boccia & Campos, 1983; Desrochers, Ricard, Décarie, & Al-lard, 1994; Feinman & Lewis, 1983; Feinman, Roberts, Hsieh, Sawyer, & Swanson, 1992).

To study social referencing, researchers have staged engaging events with ambiguous consequences and asked the infants' mothers to respond in predetermined ways, either using standardized facial expressions alone or in combination with a vocal response. The objects and events selected have included toys, strangers, animals, and a visual cliff. For example, Klinnert (1984) presented 12- and 18-month-olds with a set of novel, mobile toys and directed the infants' mothers to pose either happy, fearful, or neutral facial expressions. At both ages, the infants remained closer to the mother when she posed fear, stayed at a middle distance for neutral, and strayed farthest when she portrayed the happy expression. In a later study, Sorce, Emde, Campos, and Klinnert (1985) placed infants on a visual cliff with a 12-in. (30.48 cm) drop off. The infants' mothers posed fearful or happy expressions after the infants had been coaxed within 38.00 cm of the dropoff. Of the infants whose mothers posed happy expressions, 74% crossed over the cliff; no infant whose mother posed a fearful expression crossed.

Hornik and Gunnar (1988) introduced children 15 and 18 months old to a caged pet rabbit, their first such encounter. These infants looked to their mother and then the rabbit, making inquiring looks and vocalizations. The mothers provided af-

fective information through facial expressions and tone of voice. Initially wary infants engaged in more social referencing behaviors than the less timid infants. In another study, Hornik, Risenhooover, and Gunnar (1987) addressed the possibility that mood induction might govern these effects. Instead, they found that 12-month-olds played less with a particular toy in free play if it was the target of their mother's negative expression during a prior social-referencing phase. Effects of mothers' affect were specific rather than influential to the infants' activity in general.

Even younger infants have shown social referencing in several stranger approach studies. Ten-month-olds responded more positively to a stranger when their mothers directed positive facial, vocal, and gestural messages to their infants, although not when those same behaviors were directed to the stranger (Feinman & Lewis, 1983). Boccia and Campos (1983) reported social referencing by infants 8½ months old.

Walden and Ogan (1988) examined the development of social referencing with infants 6–9, 10–13, and 14–22 months old. They reported that all infants referred to their parent but that, beginning at 10 months old, infants' gaze became increasingly concentrated on their parent's face. In addition, younger infants looked more often and longer when their parent produced positive expressions, but the oldest infants only did so when fear was depicted. Camras and Sachs (1991) concentrated on the identity and expressiveness of the referencing target. Infants with a more expressive caretaker stayed farther from a toy, particularly in response to partially masked fear–avoidance expressions. These infants also tended to look more at fear–avoidance responses.

Of special relevance, Rosen, Adamson, and Bakeman (1992) examined “more fully how affective messages about events are conveyed by mothers to their 12-month-old infants and how this information is used” (p. 1172). Maternal messages of fear and happiness (both assigned or freely given) were effective to regulate their infant's regard for an object, whether they occurred before or after infants' initial responses to the object. Infants used social referencing often and flexibly repeated referencing to their mother, especially on the freely given fearful trials. Mothers, in turn, modulated their freely given fear displays in response to their infant. Gender-related differences were also found: Mothers presented less intense fearful messages to female infants, yet the female infants' responses were more regulated by these muted displays.

Infants of about 10 months old seem to use parental expressions to interpret an event, a process that may be influenced by a number of factors—age (Walden & Baxter, 1989; Walden & Ogan, 1988), gender (Rosen et al., 1992), temperament (Bradshaw, Goldsmith, & Campos, 1987; Feinman & Lewis, 1983; Hornik & Gunnar, 1988), cognitive ability (Desrochers et al., 1994), focus of attention (Baldwin & Moses, 1994), and referencing agent (Camras & Sachs, 1991; Hirshberg & Svejda, 1990; Klinnert, Emde, Butterfield, & Campos, 1986; Zarbata & Lamb, 1985). In conclusion, “social referencing involves an active and complex process of appraisal and judgment rather than a merely passive contagion of emotion” (Camras & Sachs, 1991, pp. 27–28).

In summary of emotion perception experiments, the data reviewed so far provide an outline for the development of infants' perception of emotional expressions during the first year. There is evidence for the detection and discrimination of facial and

vocal expressions early on. Additional data suggest that, before the second half of the first year, infants may recognize emotional expressions, as indicated by their ability to categorize facial expressions and to make intermodal matches. Late in the first year, they may use others' expressions to judge events. Before drawing firmer conclusions, however, I consider findings regarding speech perception and naturalistic interactions.

### Perception of ID Speech

Additional information about infants' perception of vocal expressions comes from investigations of their perception of ID speech. ID speech typically has a higher frequency with a wider range than AD speech. For adult listeners, ID speech facilitates communication of affect. Fernald (1989) found that adults were more accurate in categorizing speech segments with respect to communicative intent when the speech was directed to infants, for example. Fernald suggested that the exaggeration of prosodic contours in ID speech highlights the relationship between these contours and communicative intent, possibly aiding infants in detecting the meaning of vocalizations.

Fernald (1993) asserted that 5-month-olds respond differentially to ID approvals and prohibitions, even in an unfamiliar language. Infants were exposed to German, Italian, and Japanese ID speech and English AD and ID speech while looking at a photographed neutral face. There were no consistent fixation differences to the face, but observers rated infants higher on positive affect to approvals (range .27–.60, using a 3-point scale where 0 = *neutral*, 1 = *positive attention* or *tense brow*, and 2 = *smile* or *frown*) compared with prohibitions (.19–.34) for German, Italian, and English ID speech. Fernald reported that infants responded more negatively to prohibitions (.09–.19) compared with approvals (.04–.13) for the ID speech. Similarly, Papousek, Bornstein, Nuzzo, Papousek, and Symmes (1990) found that 4-month-olds looked longer to a face when fixation led to vocal ID approvals rather than ID disapprovals. Fernald concluded that ID speech is more effective than AD speech to elicit infant affect and that infants respond to the qualities of ID speech no matter what their usual language environment is. Whether infants also respond to affective information is less clear, given the overlap in positive and negative ratings to approvals and prohibitions and the overlap in negative ratings to ID prohibitions and AD speech.

Further support for the universality of the prosodic features of ID speech and its attentional and affect-communicating qualities across languages is provided by Werker, Pegg, and McLeod (1994, Cantonese) and Grieser and Kuhl (1988, Mandarin; but see Ingram, 1995). That infants are sensitive to such qualities of ID speech and can discriminate ID speech from AD speech has been demonstrated by many researchers. Kaplan and his colleagues have proposed that ID speech is more sensitizing for infants than AD speech: Four-month-olds looked longer overall at a stimulus when it was paired with ID speech (Kaplan, Goldstein, Huckeby, & Cooper, 1995). Kaplan, Jung, and Jeffers (1994) reported that ID speech served as a more effective conditioned stimulus than AD speech for facial expressions (smiling, surprise, or neutral).

Additional work has been conducted by Werker and colleagues. Werker and McLeod (1989) reported that 4.0- to 5.0-, 5.5-, and 7.5- to 9.0-month-olds preferentially looked at



videotapes of actors (both male and female) using ID speech. When the visual portion of the videotapes of a woman uttering ID or AD speech was paired with "easy listening" music instead, infants did not show a preference. Infants, especially in the youngest group, showed greater affective responsiveness in the female-ID speech conditions. Finally, unaware, untrained adults who viewed videotapes of the 4.0- to 5.5-month-olds consistently rated the infants watching a female ID speaker more positively. Werker and McLeod noted that their ID speakers were recorded while talking to a 6.0-month-old. Given that the 4.0- to 5.5-month-olds responded more to the ID displays than did the older infants, they suggested that the adults matched style of ID speech to an infant's developmental level. They did not venture to guess about how adults modify their speech, but results from Bornstein et al. (1992) may provide a clue. Bornstein et al. recorded mothers from four nations as they talked to their 5.0- and 13.0-month-old infants and analyzed these speech samples for affective and information content. Although mothers spoke more to older infants overall, a greater percentage of speech to younger infants was affective. Infants may be sensitive to relative levels of affectivity in styles of ID speech, preferring the speech that is more closely linked to their perceptual, cognitive, and social abilities.

#### Affective Expressions Presented in Social Interactions

From the laboratory studies on perception of emotion, it appears that infants may recognize bimodally presented emotional expressions by 5 to 7 months old, but researchers using other, more naturalistic methods confirmed this ability even earlier. Several of these studies have used variants of the still-face (SF) procedure; others have staged mother-infant interactions. Both fixation and behavioral responses were recorded.

#### Staged Mother-Infant Interactions

Haviland and Lelwica (1987) reported imitation and discrimination of expressions by 10-week-olds (see also Field et al., 1983). They videotaped mother-infant interactions in which mothers acted out three facial and vocal expressions noncontingently, for 20 s each. Mothers' and infants' facial behaviors were coded using the maximally discriminative facial movement coding system (MAX; Izard, 1979), mothers' tones of voice were rated by untrained coders, and event analyses (Sackett, 1979) were conducted to determine whether infants responded differentially and contingently. Haviland and Lelwica drew three major conclusions. First, by 10 weeks of age, infants responded differently to joy, anger, and sadness when the presentations were both facial and vocal. Second, infants could match or mirror expressions of joy and anger. Third, an infant's matching responses were only part of a complex but predictable set of behaviors that seem to indicate an affective response on the part of the infant.

#### Affective Reactions to Vocal Expressions

Scherer et al. (1994) reported preliminary results from studies in which researchers examined infants' reactions to violations of expectancies. During a face-to-face interaction, the experimenter's voice was filtered to produce a "metallic sound-

ing" voice associated with anger. Infants 11 to 14 months old looked more attentively on such trials, and infants from 5 to 14 months old responded with brow and mouth changes typically interpreted as surprise or interest. Some infants also showed momentary "freezing" reactions. The change in timbre was detected, as indicated by changes in their facial expressions and body movements.

#### SF Paradigm

A host of mother-infant interaction studies suggest that infants recognize affective expressions in modified interchanges. In one of the first studies, Brazelton, Koslowski, and Main (1974) altered the customary interaction of a mother and child. The mother was instructed to remain unresponsive to her 5-month-old for a part of the otherwise typical interaction. The infant looked away, looked back, grimaced, and finally withdrew from the situation. These findings have been replicated many times. Researchers typically reported that infants used facial expressions, vocalizations, and body movements in attempts to entice their mother to respond. When these efforts failed, infants reacted by turning away, frowning, and crying (Ellsworth, Muir, & Hains, 1993; Gusella, Muir, & Tronick, 1988; Mayes & Carter, 1990). Researchers have also studied more systematically what information is critical. Several aspects have been identified: type of partner (mother, stranger, object), contingency, touch, and vocal expressions. For example, Ellsworth et al. presented 3-month-olds with an unfamiliar adult and three hand puppets, varying in their resemblance to a human face. Each puppet was made to respond contingently, and all were accompanied by a melodic sound that mimicked the tonal variations of conversational speech. Although the infants looked equally to the puppets and the responsive face, they rarely smiled at a puppet. In contrast, they directed frequent grins and more vocalizations at the contingent adult. In an SF portion of the experiment, 3- and 6-month-olds also smiled more to persons than objects. They did not vary their affective responses when an object became stationary, although they increased vocalizations and grimaces when a person became unresponsive. Furthermore, as documented by others, if the mother was allowed to touch the infant during the SF period, negative reactions were reduced and smiling was increased (Gusella et al., 1988; Stack & Muir, 1990, 1992). Stack and Muir (1992) suggested that touching may be a social signal for infants, part of a set that includes a visual, auditory, and tactual mode.

Muir and Hains (1993) reviewed a series of SF studies. These studies replicate previous findings, "confirming that by 3 months-of-age infants are very sensitive to changes in adult facial expressions" (p. 182). The remainder of the review underscores the complexity of infants' responses. For example, they found that infants did not respond to a mother's interactive voice during an SF period, but a stranger's voice presented either without a face or in conjunction with an inverted face elicited positive affect. Infants did not become upset when presented maternal noncontingent behavior, but a stranger's noncontingent behavior diminished the infants' positive responses, provided that they have already interacted with the stranger in a contingent, face-to-face interaction. Muir and Hains suggested that these results emphasize the importance of context and, consequently, the need to consider task demands.

One article specifically addresses infants' responses to dynamic emotional expressions in the context of naturalistic interactions. D'Entremont (1994) conducted three studies to investigate the impact of happy, sad, and neutral facial expressions on affective displays by 4- to 6-month-olds. In the first two studies, infants looked and smiled less at an SF compared to a dynamic face, regardless of expression. In the third, when presented with a dynamically interacting face and voice, infants smiled less to the sad expression compared with the happy expression.

Izard et al. (1995) provided rare longitudinal data. Infants were observed during mother–infant interactions at 2.5, 3.0, 4.5, 6.0, 9.0, and 9.5 months. In the positive conditions, mothers expressed interest and joy facially and vocally. In the negative conditions, mothers assumed a still and silent face or expressed sadness or anger facially or vocally. The positive conditions elicited more interest and joyful expressions, and the negative conditions elicited more negative expressions. In addition, age-related increases in infant–mother expression matching occurred. Izard et al. acknowledged that it is difficult to say whether infants' expressions were "triggered simply by the perceptual coding of different maternal expressions or were related to changes in the infants' subjective states" (p. 1003), although an age-related increase in matching speaks against simple contagion.

In summary, young infants appear to discriminate affective expressions that are presented dynamically and multimodally, especially in naturalistic encounters. Their looking preferences for sound-accompanied facial expressions in laboratory studies and differential and systematic responses (especially smiling) to expressions in naturalistic settings imply recognition as well. In general, the results from the SF paradigm are consistent and point to early (3 to 4 months old) sensitivity to expressions, especially when they are experienced as dynamic, multimodal, interactive events.

### Sensory Dominance

With a step back, given results from studies of infants' discrimination of facial, vocal, and bimodally presented emotions, can one characterize such development, and does it have anything in common with other perceptual abilities? In 1930, Charlotte Bühler suggested that infants learn to discriminate emotional expressions in a specific order: first multimodal expressions, then vocal, and finally facial (Bühler & Hetzer, 1928). This is the position, with some caution, I take. From the evidence presented thus far, it looks as though infants begin to discriminate emotional expressions as early as 3 months old. Several researchers have begun to pull these data together to determine whether there is a predictable sequence with respect to modality. Three interpretations have emerged: (a) Information available to audition may be dominant, particularly when infants are presented multimodal events; (b) information given in vocal expressions is responded to earliest; and (c) young infants require both auditory and visual information for discrimination and recognition of expressions. Each of these are considered.

As described above, most studies of discrimination of vocal expressions used a procedure devised by Horowitz (1974; see also Best, McRoberts, & Sithole, 1988; Bundy, 1980; and Walker-Andrews & Grolnick, 1983). In particular, Lewkowicz (1988a, 1988b) has designed studies on sensory dominance,

modifying the original procedure by changing the auditory stimulus, the visual stimulus, or both. Based on the results of these experiments, research on the development of functional hierarchies in other mammals, and that the auditory system develops much earlier than the visual system (Gottlieb, 1971), Lewkowicz advanced a theory of early auditory dominance. To summarize, Lewkowicz visually habituated 6- and 10-month-olds to flashing checkerboards, accompanied by beeps. On the posttests, the rate of the auditory, visual, or auditory–visual information was altered. In five separate experiments (Lewkowicz, 1988a), 6-month-olds dishabituated when the rates of both the auditory and visual components were changed together. In two experiments, they also dishabituated to changes in the auditory component only. Ten-month-old infants (Lewkowicz, 1988b) dishabituated to auditory–visual changes or auditory-only changes. In four of the five experiments, they also dishabituated on posttests with visual-only changes. Lewkowicz concluded that younger infants were less proficient with visual-only changes. The overall pattern, however, indicates better discrimination for auditory–visual changes at both ages, although it is not clear whether this was an additive effect, auditory or visual information interfered with perception of its counterpart, or some other interactive effect was accountable.

These data regarding sensory dominance are limited to socially irrelevant stimuli, although Lewkowicz (1988b) originally argued that it also plays a part in the perception of social information, such as affective expressions. Along these same lines, A. J. Caron (1988; A. J. Caron et al., 1988) proposed a weaker version of auditory dominance in emotion perception. A. J. Caron et al. presented infants with videotapes of a woman posing an expression and, for some conditions, speaking in an affectively matching tone of voice (happy, sad, or angry). Four-month-olds dishabituated to a change from sad to happy for composite expressions (both facial and vocal expressions changed), provided these were shown in the sad-to-happy order. Five-month-olds dishabituated to changes in both the sad-to-happy and happy-to-sad directions for composite expressions and face-only changes. Seven-month-olds dishabituated to a change in expression in the happy-to-angry order for composites but not for facial-only displays. A. J. Caron et al. concluded that acoustic information was primary and that the pairs of emotions that they characterized as dynamically more distinct (happy and sad vs. happy and angry) were discriminated earlier.

Based on these findings and those of Walker-Andrews and Grolnick (1983), A. J. Caron (1988) proposed a developmental sequence in which infants discriminated expressions based on auditory components alone at least as early as, if not before, discrimination of composite expressions (voice and face). He suggested that, as visual resolution improves, infants detect the concordance in temporal patterns in face and voice, until finally discrimination based solely on visual information becomes possible. Fernald (1990, 1992) also argued that infants discriminate acoustic affective information first:

At the age of 5 months, when infants are not yet showing consistent selective responsiveness to positive and negative facial expressions, infants do respond differentially to positive and negative vocal expressions, suggesting that the voice is more powerful than the face as a social signal. (Fernald, 1992, p. 408)

In general, Fernald proposed that ID speech functions first as

an unconditioned stimulus to alert, soothe, and afford enjoyment or distress and then becomes effective to direct attention and modulate affect, followed by a period in which vocal and facial expressions provide "initial access to the feelings and intentions of others" (p. 403).

E. J. Gibson (1991) and Walker-Andrews and Lennon (1991) argued instead that emotion discrimination and recognition occurs first in multimodal contexts. Walker-Andrews and Lennon derived their proposal from the same results as did A. J. Caron (1988) but focused on different aspects. Specifically, they pointed to asymmetries in the data, such as order effects, and they emphasized that facial expressions always accompany vocal expressions in Horowitz's (1974) procedure. For example, in Walker-Andrews and Grolnick's (1983) study, 3-month-olds visually dishabituated only when happy vocal expressions followed sad vocal expressions. A. J. Caron et al. (1988) found order effects for sad and happy and for angry and happy expressions, but A. J. Caron interpreted dishabituation in any order as evidence for discrimination of the pair of expressions. In contrast, Walker-Andrews and Lennon suggested that fixation in these studies may reflect interest in (or detection of) particular features present at the moment rather than the result of a comparison process between the first (habituated) and second (posttest) expression. Younger infants may prefer to listen to voices characterized by higher pitch and other features characteristic of happy expressions (or ID speech). The renewal of visual attention may reflect only this initial preference. Others have made similar arguments. For example, Nelson et al. (1979; but see Ludemann & Nelson, 1988) designated order effects as examples of a "tropistic" response (much as a moth is drawn to light). Slater, Earle, Morison, and Rose (1985) have argued that, for visual patterns, "strong natural preferences . . . cannot be changed by habituating infants either to the preferred or to the nonpreferred member of a stimulus pair" (p. 52). Moreover, Malcuit, Pomerleau, and Lamarre (1988) have suggested that infants' recovery of looking time must reflect in part the arousing properties of the stimulus materials. Thelen and Smith (1994) made a strong statement about such perceptual biases in infants:

Having a bias in the system that says light is better than no light or human voices are better than auto horns does not endow the system with knowledge modules or conceptual primitives or the understanding of object properties. Rather, these are valences or tropisms similar to those exhibited by simple organisms. (p. 35)

Walker-Andrews and Lennon (1991) also pointed out that in the Horowitz procedure a vocal expression is never presented alone. Therefore, they examined discrimination of angry, happy, and sad vocal expressions, accompanied by affectively matching static facial expressions, affectively nonmatching facial expressions, or a checkerboard. Infants discriminated the vocal expression, provided a facial expression (affectively matching or nonmatching) was available. No evidence of discrimination was found when a checkerboard was present. Finally, some of the strongest evidence for infants' discriminating vocal expressiveness in ID speech (reviewed above; Fernald, 1993; Papousek et al., 1990) comes from studies in which a photograph of a face was also available.

Lewkowicz has examined the auditory dominance hypothesis with respect to dynamic stimuli (Lewkowicz, 1992) and faces

and voices (Lewkowicz, 1996; Lewkowicz & Edmondson, 1993), leading to a change in the original proposal. Lewkowicz (1992) favored a modality appropriateness hypothesis that considers the type of stimulation, specialization of the different modalities, and the fit between these two factors. In one set of experiments, Lewkowicz (1996) tested infants 4, 6, and 8 months old. In the first experiment, infants were habituated to a videotape of a woman reading a script. On posttests, the infants saw and heard the same woman reading the script (no change), a novel female face paired with the habituated female voice reading a novel script (visual only), the habituated female face paired with a novel female voice reading a novel script (auditory change only), or a change in both face and voice (auditory-visual change). Likewise, in a second experiment, Lewkowicz presented auditory-only, visual-only, and auditory-visual changes but used male and female faces and voices to highlight these changes. In these experiments, none of the infants dishabituated to a change in voice (auditory-only), but all infants dishabituated to a change in face (visual-only) or change in face and voice (auditory-visual). The third experiment contrasted auditory and visual changes but with a female face uttering ID speech or a male face uttering AD speech. Results were mixed, depending on whether infants were habituated to ID or AD speech. When infants were habituated to AD speech and tested with ID speech, 4-month-olds recovered looking to visual-only and auditory-visual changes, whereas 6- and 8-month-olds recovered to auditory-only, visual-only, and auditory-visual changes. Lewkowicz concluded that infants' responsiveness to audible and visible features of the human face can be characterized as follows: greatest recovery to bimodal changes, nearly equivalent recovery to visual-only changes, and lower discriminative recovery to auditory-only changes.

In another set of experiments, Lewkowicz and Edmondson (1993) presented 4-, 6-, and 8-month-olds with videotapes of a man or a woman talking in AD speech. Infants at all ages discriminated a change in face, a change in face and voice, or a change to an unrelated cartoon but not a change in voice only. In a second experiment, infants were presented with a man speaking in AD speech, followed by a woman singing. Infants at all ages discriminated the auditory-only, visual-only, and auditory-visual changes. In a third experiment, infants observed a woman speaking in AD speech, followed by a different woman singing. All infants dishabituated to a change in voice only, face only, and both. It is difficult to determine whether responses to the auditory and visual changes are simply additive, although there is some negative evidence. In the first experiment, the magnitude of increased fixation to a visual change was almost identical to that for the auditory-visual change. In the second and third experiments, increased fixation to the auditory-visual change was greater than that to either auditory alone or visual alone and greater to that shown to the cartoon in five of the groups (which also represents an auditory-visual change). Lewkowicz and Edmondson concluded that, when infants are presented with the type of prosody and pitch that characterize singing, they can discriminate vocal differences that are accompanied by unchanging visual information. This represents a clear shift from his original position, one foreshadowed by his prior remarks: "It is possible that the results might be different if spatially dynamic visual stimuli and patterned and socially

meaningful auditory stimuli were used” (Lewkowicz, 1988a, p. 170).

Further investigation of the roles of auditory and visual information is still required. Infants are sensitive to many acoustic parameters that must carry much of the information for affect but may not be treated as affective in isolation. Cooper (1993) has suggested that young infants respond to the quantitative (e.g., overall spectral complexity) rather than the qualitative aspects of auditory events:

It is possible that a similar type of quantitative → qualitative shift [as proposed for vision] also occurs in young infants’ auditory perception. Rather than being predisposed to attend differentially to exaggerated pitch contours, infants need to learn to extract such information from speech through experience . . . Infants may come to recognize intonation contours as salient acoustic features because of their co-occurrence with particular speakers, contexts, and other sensory stimulation (e.g., facial expressions). (p. 162)

Lewkowicz (personal communication, January 31, 1996) suggested that a number of factors—presence or lack of contingency, task demands, specific stimulus materials, and context effects—modify infants’ responsiveness.<sup>1</sup> For example, Stein, Meredith, and Wallace (1994) argued that a multimodal stimulus can have multiplicative effects. Such effects cannot be anticipated from those of the single modality components. Some effects may be specific: In the Lewkowicz (1996; Lewkowicz & Edmondson, 1993) studies, infants responded differently, depending on whether AD or ID speech was presented during the habituation sequence. Other effects may be more subtle or indirect: The checkerboard visual stimulus used by Walker-Andrews and Lennon (1991) might have simply distracted infants, obscuring their ability to discriminate. The emergence of responsiveness to auditory and visual information that characterizes emotional expressions may vary widely, depending on specifics of the information as well as the developmental status of the infant.

### Development of the Perception of Speech Sounds

A major proposal in this article is that the development of emotion perception parallels the development of perception in other domains. In the following section, I discuss some aspects of speech perception (speech sounds) and its development to determine whether this proposal can be supported. Arguably, speech perception is a select example because it is also concerned with the perception of persons (Neisser, 1994). However, the development of competencies in emotion perception and speech perception is important, protracted, and likely to be challenging because both types of information are multifaceted and complicated.

In 1971, Eimas, Siqueland, Jusczyk and Vigorito presented astounding results. They demonstrated that 1-month-olds could discriminate English phonemes and, moreover, that the distinction was categorical. Many have replicated these original results (e.g., Aslin, Pisoni, & Jusczyk, 1983; Eimas, 1975; Trehub, 1976; Werker & Tees, 1984). Others have shown that infants can show “perceptual constancy” for speech sound categories: For example, infants treat a vowel sound the same across different speakers (Kuhl, 1979).

Superficially, speech perception would seem to be a purely

auditory process. Meltzoff and Kuhl (1994) argued, however, that infants “code faces and speech as intermodal objects of perception” (p. 335). They based their argument on data that reveal that vision (Dodd & Campbell, 1987; Grant, Ardell, Kuhl, & Sparks, 1985; Green & Kuhl, 1991; Green, Kuhl, Meltzoff, & Stevens, 1991; Massaro, 1987, 1994; McGurk & MacDonald, 1976; Summerfield, 1979, 1987) and touch (Grant, Ardell, Kuhl, & Sparks, 1986; Sparks, Ardell, Bourgeois, Wiedmer, & Kuhl, 1979) both play a role in what a participant reports hearing. Some evidence for this is anecdotal; other evidence is empirical. One feels as if one hears “better” when one can view a speaker’s mouth movements. Sumbly and Pollack (1954) reported that to watch a speaker’s lip is tantamount to a gain of about 20 dB SPL in intensity. To look at a speaker’s face while one listens to a pure-tone signal that parallels the pitch and amplitude of the voice leads to 80% intelligibility, compared with 0% for a tone alone and 37% for a face alone (Grant et al., 1985).

Additional evidence comes from experiments on an auditory–visual illusion, the McGurk effect (McGurk & MacDonald, 1976). When auditory information for /b/ was combined with visual information for /g/, perceivers reported hearing /d/. Green and Kuhl (1989, 1991) suggested that integration of the two sources of information takes place before the listener categorizes the syllable, before “the speech stream . . . has been rigidly coded as having a defined and specific place or manner of articulation” (Meltzoff & Kuhl, 1994, p. 351). Further support for this conjecture is provided by Green et al. (1991), who paired a male face with a female voice and vice versa. Although the participants noticed the discrepancy, they also reported the usual number of auditory–visual illusions. Analogous results were found for tactual influences. Grant et al. (1986) reported that pure-tone information, matched to the pitch of a speaker’s voice and delivered by way of electrodes placed on the skin, combined with lip reading increased intelligibility 20% over lip reading alone.

Young infants are also sensitive to correspondences in speech presented to their eye and ear. Four-month-olds can detect a match between a face and voice that produced the same vowel sound (Kuhl & Meltzoff, 1984, 1988; MacKain, Studdert-Kennedy, Spieker, & Stern, 1983). When a 3-month-old hears a vowel that matches or mismatches the sound an adult is visibly but silently producing, the infant selectively imitates the sound that is presented bimodally (Kuhl & Meltzoff, 1982; Legerstee, 1990). Unlike adults, however, infants did not match pure tones or three-tone nonspeech analogs to visually presented vowels, indicating that, for infants, detection of cross-modal correspondences for speech requires the whole speech stimulus. As suggested for emotion, infants may be responsive first to wholes or gestalts (Lewkowicz, 1996; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991), which later may be differentiated into component features.

Meltzoff and Kuhl (1994) also discussed the role of visual, auditory, and proprioceptive information in production. Data obtained from speech produced by children with hearing impairments, children with visual impairments, and controls highlight some of the differences. According to Meltzoff and Kuhl, the

<sup>1</sup> A reviewer also suggested this.

babbling of infants with hearing impairments did not duplicate the pattern universal among hearing infants. The timing of specific advances was different, the duration of utterances was different, and the phonetic content was unusual. Infants with hearing impairments produced disproportionately more bilabial sounds; whereas control children produced a higher proportion of "hidden" sounds, such as /g/. In addition, blind children learned sounds with a visible articulation pattern more slowly (Mills, 1987). Meltzoff and Kuhl (1994) concluded that

infants who are engaged in cooing and babbling in their bassinets are engaged in serious business: They are mastering quite general rules about the auditory consequences of their own vocal tract manipulations. They are solidifying an auditory-articulatory intermodal map of speech. In developing this map they use auditory and proprioceptive information from the self and visual information from others to learn what to do with their own vocal tracts when producing speech. (p. 358)

### Differentiation of Information for Expressions

What can I conclude about infants' perception of emotional expressions? First, the data on infants' perception of facial and vocal expressions have many shortcomings. Most of the experiments have been cross-sectional in design, and the stimulus materials and tasks have varied widely. Many researchers have used standard photographs; others have used videotapes designed as representing some particular affect by untrained judges or by using facial coding systems. Similarly, voice tokens have been either labeled by untrained judges or selected to fit criteria set by Scherer (1986). This variability leads to difficulties in interpretation, but this very variability also allows me to conclude that infants are sensitive to information for affect and that they discriminate emotional expressions across a wide array of exemplars, which sample the full range of emotional expressions.

The data reviewed above suggest that very young infants are sensitive to facial and vocal information that potentially specifies emotional expressions. By way of a summary, infants can detect suprasegmental information in their mother's voices that must have been available in utero, thus showing a preference to listen to those sounds. Early on, infants preferentially look to faces that differ in features, such as toothiness. Such feature differences allow them to discriminate static expressions of discrete emotions. Data regarding the ability to use features to discriminate facial expressions and vocal expressions indicate that infants can tell such expressions apart in the first few months of life. By combining data from studies of generalization of facial expressions, studies documenting intermodal matching, and results from studies on infants' responses to social interactions, I can conclude that infants may, depending on one's criteria, recognize emotional expressions by the second half of the first year. The issue remains, however, as to how these abilities develop.

### *Modularity as an Explanation*

Although in no published article has a researcher applied a modularity perspective to infants' emotion perception, in principle such a perspective has much to offer. Baron-Cohen's (1994) model, for example, could be extended to the data. First, the

infant would have a set of special purpose modules specifically tuned to emotional information. Each would switch on at varied but specific ages. Some modules might respond to auditory information only, some to visual information only, and some to auditory-visual information. Development, in this view, would follow the maturation of the autonomous, intermodal, or modality-specific encapsulated modules. For example, an early maturing module might be sensitive to frequency, contributing to the preferences infants show for ID speech. A separate module might respond to facial features. Later maturing modules might respond to affectively relevant information, multimodal information, or both. The extant cross-sectional data could presumably be explained by a set of such modules. The unevenness in age-related findings could be explained away by references to sampling, methodological differences, and so forth. Longitudinal studies are necessary to fully examine this claim.

However, this model could not easily handle findings that point to context effects. For example, Lewkowicz (1996, 1997) modified the auditory-dominance stance in light of two sets of findings: (a) It did not describe the infants' discrimination of dynamic and socially relevant stimuli, and (b) infants treated types of speech differently. Infants were relatively unaffected by audible differences in AD speech as late as 8 months old, yet contrasts provided by singing, ID speech, or vocal expressions were detected and discriminated months earlier. Such findings argue for "deep correspondences in the heterogeneous systems that make up mind" (Thelen & Smith, 1994, p. 36). Modularity approaches do not capture the dynamic nature of the interactions that occur throughout the development of emotion perception.

### *My Proposal*

My proposal does not depend on a set of special mechanisms but builds on what is known about infants' perceptual development in general (E. J. Gibson & Spelke, 1983; Muir, Humphrey, et al., 1994). Infants appear to experience a world of perceptual unity (e.g., Bahrick, 1983, 1988; Dodd, 1979; E. J. Gibson & Walker, 1984; Meltzoff & Borton, 1979; Mendelson & Ferland, 1982; Spelke & Owsley, 1979; Walker, 1982; Walker-Andrews & Lennon, 1985; but see Lewkowicz, 1991; Maurer, 1993; and Piaget, 1952). To learn about multimodal specification of events is an early achievement in the perceptual development of an infant, with such abilities present at most ages tested (for a review, see Rose & Ruff, 1987), although there is an ongoing debate with respect to whether the specific development of intermodal perception follows a course of differentiation or integration. Given this early sensitivity, my proposal is that infants may first recognize the affective expressions of others as part of a unified multimodal event that has a unique communicative affordance. Faces and voices are typically experienced together, as part of an event that also includes touch and smell, although research focuses on auditory and visual information. As noted by Flavell (1985),

in the extralaboratory world people do not present themselves to babies as voiceless faces or faceless voices . . . . Moreover, the face and the voice are unified in space and time: The voice and the face share the same spatial location, and the face's mouth movements are temporally synchronized. In addition, certain specific faces always co-occur with certain specific voices . . . . Finally,

how each face looks and acts on a given occasion is highly correlated with how its voice sounds; for instance, happy and sad voices usually accompany happy and sad faces, respectively. (p. 133)

The infant learns to differentiate these two modes (auditory and visual) of specification (cf. Maurer, 1993) and detect abstract invariants that specify the same emotional meaning (Walker-Andrews, 1988).

In this view, young infants detect unimodal information that potentially specifies the meaning of an expression. They detect acoustic parameters, such as timbre and frequency, that provide information for affect. Even the neonate can detect the fundamental frequencies of two different voices and discriminate them (DeCasper & Spence, 1986), but there is no evidence that this information specifies an emotional nuance to the infant. Similarly, feature differences in facial expressions (lift of the brow or presence of teeth) can be detected several months postnatally (R. F. Caron, Caron, & Myers, 1982), after the infant's visual acuity has improved enough to detect static information—it is even earlier for dynamic information (R. F. Caron et al., 1985; D'Entremont, 1994). As a consequence, older infants may discriminate a face with a toothy grin from one with a tight-lipped grimace or discriminate a high-pitched voice from a low-pitched one. This modality-specific information, however, is not appreciated as affective (allowing for recognition) until somewhat later.

Recognition, it is proposed, first occurs in multimodal contexts. The critical information specifying an emotion is found in the overall dynamic flow, particularly in the invariant patterns of movement and change undergone by facial musculature, body, and voice (Ekman & Friesen, 1978; Fogel, 1993). This does not mean that emotional information resides only in multimodal portrayals, but for infants dynamic, naturalistic, and multimodal presentations may be the optimal stimuli. An infants' auditory, visual, and haptic systems are well coordinated even at birth, so infants can attend to intermodal correspondences very early. For example, the infants' exploratory head and eye movements may be elicited by a human voice in the first few days (Butterworth & Castillo, 1976). By 3–5 months old, infants begin to discriminate bimodally presented faces (Burnham, 1993) and affective expressions from one another (Gusella et al., 1988; Walker-Andrews & Lennon, 1991), if not earlier (Field et al., 1983), especially in familiar contexts (Montague, 1995) or with familiar persons (Haviland & Lelwica, 1987).

There are a wealth of intermodal relationships provided by simultaneously available acoustic, visual, and tactual presentations of emotion to which infants may be sensitive. Although some expressions are more easily discriminated (A. J. Caron et al., 1988), perhaps because of the infants' experiences or properties of the expressions themselves, progress is rapid after infants begin to make such discriminations and recognize these expressions as emotional. At about the same time as infants are able to discriminate facial expressions from one another and vocal expressions from one another, they are able to respond differentially to multimodally presented expressions in naturalistic settings. Recognition of modality-specific expressions follows apace. Along the ideas put forth by Bahrick (1992, 1994), if infants perceive a bimodally presented expression as a single event, then an opportunity to explore other properties of the event, including modality-specific properties, is provided. In other words, infants become able to differentiate the structure

of the complex event, attending to and learning about the presence of modality-specific information in that event. Both Thelen (1986; Fogel & Thelen, 1987) and E. J. Gibson (1988) have pointed out how new achievements emerge through the maturation and reorganization of systems. In this case, as infants begin to differentiate and recognize multimodally presented emotional expressions, they return to and discriminate modality-specific information in a new way. Moreover, as the proficiency of each modality improves, other means to examine the expressions becomes less redundant; the sensory systems are used to supplement one another with respect to modality-specific properties (Muir, Humphrey, et al., 1994). Differences in pitch now denote affective expression, as indicated by 5-month-old infants' unsuccessful discrimination of vocal expressions in the presence of a checkerboard (Walker-Andrews & Lennon, 1991) but successful discrimination when any facial expression is present.

This hypothesis is not entirely new but reflects others' thinking about the perception of emotion and its development. For example, Nelson (1987) suggested that young infants are capable of discriminating facial expressions on the basis of some of the facial features that distinguish different emotions but that "the perception of facial expressions as expressions does not develop until the second half of the first year" (Nelson, 1985, p. 111). Bühler (1930) and Charlesworth and Kreutzer (1973) came to similar conclusions long before. Klinnert et al. (1983) posed developmental levels on which infants perceive facial expressions: no discrimination (0–6 weeks), discrimination devoid of understanding (6 weeks–5 months), emotional resonance in which infants directly experience another's emotion (5–9 months), and social referencing in which infants use the motions of others to guide their own actions (9+ months).

Others have also written of the intermodal nature of infants' perception as it relates to faces and voices. Werker and MacLeod (1989) emphasized the intermodal aspects of infants' perception of ID speech: They concluded that the "most effective stimulus [for ID speech] will likely be found to be the multi-modal face plus voice normally experienced by the infant" (p. 243). Meltzoff and Moore (1993) proposed a much more radical view: They suggested that "faces are special and meaningful to infants in part because infants experience their own faces through proprioception. The visual pattern provided by a face can be assimilated to infants' own felt experiences" (p. 211). The face is, according to Meltzoff and Moore, a cross-modal stimulus with great social significance to the neonate. They also proposed that a supramodal network unites one's own and others' bodily actions into a common framework. Similarly, Turkewitz (1993) took a developmental perspective that emphasizes how sensory systems develop and are deployed in new ways, leading to reorganizations in domains such as emotion perception. He suggested that, at the same time that faces are likely to be "salient by virtue of their being presented in conjunction with multimodal stimulation consequent upon being held, rocked, and spoken or sung to" (p. 137), visual details, because of limited acuity, are indistinct and blurry. Infants also respond to the expressions of others with vocal and facial expressions of their own, allowing for kinesthetic information and detection of contingency. Turkewitz (personal communication, April 19, 1995) asserted that, even before birth, infants may be aware of maternal physiological changes (including those that accompany emotional expressions). In his view, if these changes are detected

and experienced together with auditory stimulation provided by maternal vocal expressions, then the way is laid for paired association of these sources of information in utero. How infants' own emotional responses interact with their perception of others' expressions is an area well worth exploring (Stern, 1985).

As suggested by these researchers, with development infants come to perceive the affordances of emotional expressions, moving away from a reliance on multimodal stimulation. But for all people, the detection of meaning is less difficult in more naturalistic settings. What is necessary to perceive the affordances of others, their emotions, and intentions is social interactions, not a sense of self or an ability to construct meaning out of an incomprehensible stimulus array. To perceive other persons is no different from objects and events: People perceive the affordances of others by observing their expressions, actions, and physical properties.

Veridical information about people and their interactions is available in dynamic, ongoing stimulus events . . . Perception itself is conceptualized as a dynamic process in which an active perceiver comes to recognize the potential of the environment through exploration and behavior. (Berry & Misovich, 1994, p. 139)

The infant learns about others and their affective expressions when given ample opportunities to look, listen, and participate in social interactions. The perceptual systems of infants seem attuned to invariants that specify social events early on; with development, there is an increasing differentiation of the information for affect.

## References

- Aldridge, M. (1994, June). *Newborns' perception of emotion in voices*. Paper presented at the International Conference on Infant Studies, Paris, France.
- Andrew, R. J. (1963, November 22). Evolution of facial expression. *Science*, *142*, 1034–1041.
- Aslin, R. N., Pisoni, D. B., & Jusczyk, P. W. (1983). Auditory development and speech perception in infancy. In M. M. Haith & J. J. Campos (Eds.), *Infancy and developmental psychobiology* (pp. 573–687). New York: Wiley.
- Bahrack, L. E. (1983). Infants' perception of substance and temporal synchrony. *Infant Behavior and Development*, *6*, 429–450.
- Bahrack, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development*, *59*, 197–209.
- Bahrack, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology*, *53*, 180–199.
- Bahrack, L. E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology*, *6*, 111–123.
- Baldwin, D. A., & Moses, L. J. (1994). Early understanding of referential intent and attentional focus: Evidence from language and emotion. In M. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 133–150). Hillsdale, NJ: Erlbaum.
- Banks, M. S. (1980). The development of visual accommodation during early infancy. *Child Development*, *51*, 646–666.
- Banks, M., & Ginsberg, A. P. (1985). Infant visual preferences: A review and new theoretical treatment. In H. W. Reese (Ed.), *Advances in child development and behavior* (pp. 207–246). New York: Academic Press.
- Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive*, *13*, 513–552.
- Barrera, M. E., & Maurer, D. (1981). The perception of facial expressions by the three-month-old. *Child Development*, *52*, 203–206.
- Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*, 2049–2058.
- Berry, D. S., & Misovich, S. J. (1994). Methodological approaches to the study of social event perception. *Personality and Social Psychology Bulletin*, *20*, 139–152.
- Bertenthal, B. I., Proffitt, D. R., Spetner, N. B., & Thomas, M. A. (1985). The development of infant sensitivity to biomechanical motions. *Child Development*, *56*, 531–543.
- Best, C. T., McRoberts, G. W., & Sithole, N. N. (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 345–360.
- Boccia, M., & Campos, J. J. (1983, April). *Maternal emotional signaling: Its effects on infants' reaction to strangers*. Paper presented at the Society for Research in Child Development meeting, Detroit, MI.
- Bornstein, M. H., Tal, J., Rahn, C., Galperin, C. Z., Pecheux, M., Lamour, M., Toda, S., Azuma, H., Ogino, M., & Tamis, L. C. (1992). Functional analysis of the contents of maternal speech to infants of 5 and 13 months in four cultures: Argentina, France, Japan, and the United States. *Developmental Psychology*, *28*, 593–603.
- Bower, T. G. R. (1966, February 18). Slant perception and shape constancy in infants. *Science*, *151*, 832–834.
- Bradshaw, D. L., Goldsmith, H. H., & Campos, J. J. (1987). Attachment, temperament, and social referencing. *Infant Behavior and Development*, *10*, 223–231.
- Brazelton, T. B., Koslowski, B., & Main, W. (1974). The origins of reciprocity: The early mother–infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). New York: Wiley.
- Bruyer, R., Laterre, C., Seron, X., Feyerstein, P., Strypstein, E., Pierrard, E., & Rectem, D. (1983). A case of prosopagnosia with some preserved covert remembrance of familiar faces. *Brain and Cognition*, *2*, 257–284.
- Bühler, C. (1930). *The first year of life*. New York: Day.
- Bühler, C., & Hetzer, H. (1928). Das erste verstandnis für ausdruck im ersten lebensjahr [The early understanding of expressions in the first year of life]. *Zeitschrift für Psychologie*, *107*, 50–61.
- Bundy, R. S. (1980). Discrimination of sound localization cues in young infants. *Child Development*, *51*, 292–294.
- Burnham, D. (1993). Visual recognition of mother by young infant: Facilitation by speech. *Perception*, *22*, 1133–1153.
- Bushnell, I. W. R. (1979). Modification of the externality effect in young infants. *Journal of Experimental Child Psychology*, *28*, 211–229.
- Bushnell, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*, 3–15.
- Butterfield, E. C., & Siperstein, G. N. (1970). Influence of contingent auditory stimulation upon non-nutritional suckle. In J. Bosma (Ed.), *Third Symposium on Oral Sensation and Perception: The mouth of the infant* (pp. 313–334). Springfield, IL: Charles C Thomas.
- Butterworth, G., & Castillo, M. (1976). Coordination of auditory and visual space in newborn human infants. *Perception*, *5*, 155–160.
- Campos, J. J., Barrett, K. C., Lamb, M. E., Goldsmith, H. H., & Stenberg, C. (1983). Socioemotional development. In P. H. Mussen (Ed.), *Infancy and developmental psychobiology* (pp. 783–915). New York: Wiley.
- Camras, L. A., & Sachs, V. B. (1991). Social referencing and caretaker expressive behavior in a day care setting. *Infant Behavior and Development*, *14*, 27–36.
- Caron, A. J. (1988, April). *The role of face and voice in infant discrimi-*

- nation of naturalistic emotional expressions. Paper presented at the International Conference on Infant Studies, Washington, DC.
- Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child Development, 59*, 604–616.
- Caron, R. F., Caron, A. J., & Myers, R. S. (1982). Abstraction of invariant face expressions in infancy. *Child Development, 53*, 1009–1015.
- Caron, R. F., Caron, A. J., & Myers, R. S. (1985). Do infants see emotional expressions in static faces? *Child Development, 56*, 1552–1560.
- Charlesworth, W. R., & Kreutzer, M. A. (1973). Facial expressions of infants and children. In P. Ekman (Ed.), *Darwin and facial expression* (pp. 91–167). New York: Academic Press.
- Chevalier-Skolnikoff, S. (1973). Facial expression of emotion in non-human primates. In P. Ekman (Ed.), *Darwin and facial expression* (pp. 11–89). New York: Academic Press.
- Cooper, R. P. (1993). The effect of prosody on young infants' speech perception. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (pp. 137–167). Norwood, NJ: Ablex.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595.
- Culp, R. E., & Boyd, E. F. (1974). Visual fixation and the effect of voice quality and content differences in 2-month-old infants. *Monographs of the Society for Research in Child Development, 39*(5–6, Serial No. 158).
- Cummings, E. M. (1995). Security, emotionality, and parental depression: A commentary. *Developmental Psychology, 31*, 425–427.
- Dawson, G. (1994). Frontal electroencephalographic correlates of individual differences in emotion expression in infants: A brain perspective on emotion. *Monographs of the Society for Research in Child Development, 59*(2–3, Serial No. 240), 135–151.
- DeCasper, A. J., & Fifer, W. P. (1980, June 6). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*, 1174–1176.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*, 133–150.
- D'Entremont, B. (1994, June). *Young infants' responding to static and dynamic happy and sad expressions during a social interaction*. Paper presented at the International Conference on Infant Studies, Paris, France.
- Desrochers, S., Ricard, M., Décarie, T. G., & Allard, L. (1994). Developmental synchrony between social referencing and Piagetian sensorimotor causality. *Infant Behavior and Development, 17*, 303–309.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review, 3*, 357–370.
- Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology, 11*, 478–484.
- Dodd, B., & Campbell, R. (Eds.). (1987). *Hearing by eye: The psychology of lip-reading*. London: Erlbaum.
- Ecklund-Flores, L., & Turkewitz, G. (1996). Asymmetric headturning to speech and nonspeech in human newborns. *Developmental Psychology, 29*, 205–217.
- Eibl-Eibesfeldt, I. (1970). *Ethology: The biology of behaviour*. New York: Holt, Rinehart, & Winston.
- Eimas, P. D. (1975). Speech perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (pp. 193–231). New York: Academic Press.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971, January 22). Speech perception in infants. *Science, 171*, 303–306.
- Ekman, P., & Friesen, W. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
- Ellsworth, C. P., Muir, D. W., & Hains, S. M. J. (1993). Social competence and person-object differentiation: An analysis of the still-face effect. *Developmental Psychology, 29*, 63–73.
- Everhart, V., & Henry, S. (1992, May). *Indicators of early empathy: Newborns learning the meaning of changes in facial expression of emotion within 5 minutes*. Paper presented at the International Conference on Infant Studies, Miami, FL.
- Feinman, S., & Lewis, M. (1983). Social referencing at ten months: A second-order effect on infants' responses to strangers. *Child Development, 54*, 878–887.
- Feinman, S., Roberts, D., Hsieh, K.-F., Sawyer, D., & Swanson, D. (1992). A critical review of social referencing in infancy. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 15–55). New York: Plenum.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497–1510.
- Fernald, A. (1990, March). *Themes and variations: Cross-cultural comparisons of melodies in mothers' speech*. Paper presented at the Seventh International Conference on Infant Studies, Montreal, Quebec, Canada.
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford, England: Oxford University Press.
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*, 657–674.
- Field, T. (1994). The effects of mother's physical and emotional unavailability on emotion regulation. *Monographs of the Society for Research in Child Development [Special issue], 59*(2–3, Serial No. 240), 208–227.
- Field, T., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior and Development, 7*, 19–26.
- Field, T. M., Woodson, R., Cohen, D., Greenberg, R., Garcia, R., & Collins, K. (1983). Discrimination and imitation of facial expressions by term and preterm neonates. *Infant Behavior and Development, 6*, 485–490.
- Field, T., Woodson, R., Greenberg, R., & Cohen, D. (1982, October 8). Discrimination and imitation of facial expressions by neonates. *Science, 218*, 179–181.
- Flavell, J. H. (1985). *Cognitive development* (2nd ed.). New York: Prentice-Hall.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fogel, A. (1993). Two principles of communication: Co-regulation and framing. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicative development* (pp. 9–22). New York: Routledge.
- Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology, 23*, 747–761.
- Fox, R., & McDaniel, C. (1982, October 29). The perception of biological motion by human infants. *Science, 218*, 486–487.
- Fridlund, A. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60*, 229–240.
- Garner, W. R. (1981). The analysis of unanalyzed perceptions. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 119–139). Hillsdale, NJ: Erlbaum.
- Gelfand, D. M., & Teti, D. M. (1990). The effects of maternal depression on children. *Clinical Psychology Review, 10*, 329–359.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
- Gibson, E. J. (1983). Development of knowledge about intermodal unity: Two views. In L. S. Liben (Ed.), *Piaget and the foundations of knowledge* (pp. 19–41). Hillsdale, NJ: Erlbaum.
- Gibson, E. J. (1988). Exploratory behavior in the development of per-



- ceiving, acting, and the acquiring of knowledge. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (pp. 1–41). Palo Alto, CA: Annual Reviews.
- Gibson, E. J. (1991). *An odyssey in learning and perception*. Cambridge, MA: MIT Press.
- Gibson, E. J., & Spelke, E. S. (1983). The development of perception. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Cognitive development* (pp. 1–76). New York: Wiley.
- Gibson, E. J., & Walker, A. S. (1984). Development of knowledge of visual-tactile affordances of substance. *Child Development*, *55*, 453–460.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum. (Reprinted 1986)
- Goren, C. C., Sarty, M., & Wu, P. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*, 544–549.
- Gottlieb, G. (1971). Ontogenesis of sensory function in birds and mammals. In E. Tobach, L. R. Aronson, & E. Shaw (Eds.), *The biopsychology of development* (pp. 67–128). New York: Academic Press.
- Grant, K. W., Ardell, L. H., Kuhl, P. K., & Sparks, D. W. (1985). The contribution of fundamental frequency, amplitude envelope, and voice duration cues to speechreading in normal-hearing subjects. *Journal of the Acoustical Society of America*, *77*, 671–677.
- Grant, K. W., Ardell, L. A. H., Kuhl, P. K., & Sparks, D. W. (1986). The transmission of prosodic information via an electrotactile speechreading aid. *Ear and Hearing*, *7*, 328–335.
- Green, K. P., & Kuhl, P. K. (1989). The role of visual information in the processing of place and manner features in speech perception. *Perception and Psychophysics*, *45*, 34–42.
- Green, K. P., & Kuhl, P. K. (1991). Integral processing of visual place and auditory voicing information during phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 278–288.
- Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, genders, and sensory modality: Female faces and male voices in the McGurk effect. *Perception and Psychophysics*, *50*, 524–536.
- Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, *24*, 14–20.
- Gusella, J. L., Muir, D. W., & Tronick, E. Z. (1988). The effect of manipulating maternal behavior during an interaction on 3- and 6-month-olds' affect and attention. *Child Development*, *59*, 1111–1124.
- Gwiazda, J., Bauer, J., & Held, R. (1989). From visual acuity to hyperacuity: A 10-year update. *Canadian Journal of Psychology*, *43*, 109–120.
- Habib, M. (1986). Visual hypoemotionality and prosopagnosia associated with right temporal lobe isolation. *Neuropsychologia*, *24*, 577–582.
- Hainline, L., & Abramov, I. (1992). Assessing visual development: Is infant vision good enough? In C. K. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (pp. 39–102). Norwood, NJ: Ablex.
- Haith, M. M., Bergman, T., & Moore, M. J. (1977, November 25). Eye contact and face scanning in early infancy. *Science*, *198*, 853–855.
- Haviland, J. M., & Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Developmental Psychology*, *23*, 97–104.
- Hirshberg, L. M., & Svejda, M. (1990). When infants look to their parents: I. Infants' social referencing of mothers compared to fathers. *Child Development*, *61*, 1175–1186.
- Hobson, R. P. (1989). Beyond cognition: A theory of autism. In G. Dawson (Ed.), *Autism: New perspectives on diagnosis, nature and treatment* (pp. 22–48). New York: Guilford Press.
- Hobson, R. P., Ouston, J., & Lee, T. (1988). Emotion recognition in autism: Coordinating faces and voices. *Psychological Medicine*, *18*, 911–923.
- Hornik, R., & Gunnar, M. R. (1988). A descriptive analysis of infant social referencing. *Child Development*, *59*, 626–634.
- Hornik, R., Risenhoover, N., & Gunnar, M. R. (1987). The effects of maternal positive, neutral, and negative affective communication on infant responses to new toys. *Child Development*, *58*, 937–944.
- Horowitz, F. D. (1974). Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development*, *39*(5–6, Serial No. 158).
- Hutt, S. J., Hutt, C., Leonard, H. G., von Bermuth, H., & Muntjewerff, W. F. (1968, June 1). Auditory responsivity in the human neonate. *Nature*, *218*, 888–890.
- Ingram, D. (1995). The cultural basis of prosodic modifications to infants and children: A response to Fernald's universalist theory. *Journal of Child Language*, *22*, 223–233.
- Izard, C. E. (1979). *The maximally discriminative facial movement coding system (MAX)*. Newark: University of Delaware Press.
- Izard, C. E., Fantauzzo, C. A., Castle, J. M., Haynes, O. M., Rayias, M. R., & Putnam, P. H. (1995). The ontogeny and significance of infants' facial expressions in the first 9 months of life. *Developmental Psychology*, *31*, 997–1013.
- Johnson, M. H., Dziurawiec, S., Bartrip, J., & Morton, J. (1992). The effects of movement of internal features on infants' preferences for face-like stimuli. *Infant Behavior and Development*, *15*, 129–136.
- Johnson, M. H., Dziurawiec, S., Ellis, H. D., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*, 1–21.
- Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development*. Cambridge, England: Basil Blackwell.
- Kaplan, P. S., Goldstein, M. H., Huckleby, E. R., & Cooper, R. P. (1995). Habituation, sensitization, and infants' responses to motherese speech. *Developmental Psychobiology*, *28*, 45–57.
- Kaplan, P., Jung, P., & Jeffers, C. (1994, June). *Differential effects of infant-directed versus adult-directed speech as signals for adult female faces*. Paper presented at the International Conference on Infant Studies, Paris, France.
- Kaufmann, F., Stucki, M., & Kaufmann-Hayoz, R. (1985). Development of infants' sensitivity for slow and rapid motions. *Infant Behavior and Development*, *8*, 89–98.
- Kaufmann, R., & Kaufmann, F. (1980). The face scheme in 3- and 4-month-old infants: The role of dynamic properties of the face. *Infant Behavior and Development*, *3*, 331–339.
- Kestenbaum, R., & Nelson, C. A. (1990). The recognition and categorization of upright and inverted expressions by 7-month-old infants. *Infant Behavior and Development*, *13*, 497–511.
- Klinnert, M. (1984). The regulation of infant behavior by maternal facial expression. *Infant Behavior and Development*, *7*, 447–465.
- Klinnert, M., Campos, J. J., Sorce, J., Emde, R. N., & Svejda, M. (1983). Emotions as behavior regulators: Social referencing in infancy. In R. Plutchik & H. Kellerman (Eds.), *Emotions in early development* (pp. 57–86). New York: Academic Press.
- Klinnert, M. D., Emde, R. N., Butterfield, P., & Campos, J. J. (1986). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. *Developmental Psychology*, *22*, 427–432.
- Köhler, W. (1940). *Dynamics in psychology*. New York: Liveright.
- Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, *37*, 1539–1553.
- Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child Development*, *57*, 1054–1061.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66*, 1668–1679.
- Kuhl, P. K., & Meltzoff, A. N. (1982, December 10). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141.

- Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development*, 7, 361–381.
- Kuhl, P. K., & Meltzoff, A. N. (1988). Speech as an intermodal object of perception. In A. Yonas (Ed.), *Perceptual development in infancy* (pp. 235–266). Hillsdale, NJ: Erlbaum.
- LaBarbera, J. D., Izard, C. E., Vietze, P., & Parisi, S. A. (1976). Four- and six-month-old infants visual responses to joy, anger, and neutral expressions. *Child Development*, 47, 535–538.
- Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior and Development*, 13, 343–354.
- Leslie, A. M. (1987). Pretense and representation: The origins of “theory of mind.” *Psychological Review*, 94, 412–426.
- Leslie, A. M. (1991). The theory of mind impairment in autism: Evidence for a modular mechanism of development? In A. Whiten (Ed.), *Natural theories of mind* (pp. 63–78). Oxford, England: Basil Blackwell.
- Leslie, A. M. (1994a). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211–238.
- Leslie, A. M. (1994b). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. A. Hirshfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). New York: Cambridge University Press.
- Lewkowicz, D. J. (1988a). Sensory dominance in infants. 1: Six-month-old infants’ response to auditory–visual compounds. *Developmental Psychology*, 24, 155–171.
- Lewkowicz, D. J. (1988b). Sensory dominance in infants. 2: Ten-month-old infants’ response to auditory–visual compounds. *Developmental Psychology*, 24, 172–182.
- Lewkowicz, D. J. (1991). Development of intersensory functions in human infancy: Auditory/visual interactions. In M. J. Weiss & P. R. Zelazo (Eds.), *Newborn attention* (pp. 308–338). Norwood, NJ: Ablex.
- Lewkowicz, D. J. (1992). Infants’ responsiveness to the auditory and visual attributes of a sounding/moving stimulus. *Perception and Psychophysics*, 52, 519–528.
- Lewkowicz, D. J. (1996). Infants’ response to the audible and visible properties of the human face. I: Role of lexical/syntactic content, temporal synchrony, gender, and manner of speech. *Developmental Psychology*, 32, 347–366.
- Lewkowicz, D. J. (1997). *Infants’ response to the audible and visible properties of the human face. II: Effects of singing*. Manuscript submitted for publication.
- Lewkowicz, D. J., & Edmondson, D. (1993, March). *Infants’ responsiveness to the multimodal properties of the human face*. Presented at the Society for Research in Child Development meeting, New Orleans, LA.
- Ludemann, P. M. (1991). Generalized discrimination of positive facial expressions by seven- and ten-month-old infants. *Child Development*, 62, 55–67.
- Ludemann, P., & Nelson, C. (1988). Categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology*, 24, 492–501.
- MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983, March 18). Infant intermodal speech perception is a left-hemisphere function. *Science*, 219, 1347–1349.
- Malatesta, C. Z., Culver, C., Teasman, J. R., & Shepard, B. (1988). The development of emotion expression during the first two years of life. *Monographs of the Society for Research in Child Development*, 54(1–2, Serial No. 219).
- Malcuit, G., Pomerleau, A., & Lamarre, G. (1988). Habituation, visual fixation and cognitive activity in infants: A critical analysis and attempt at a new formulation. *Cahiers de Psychologie Cognitive*, 8, 415–441.
- Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.
- Massaro, D. W. (1994). Bimodal speech perception across the life span. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 371–399). Hillsdale, NJ: Erlbaum.
- Maurer, D. (1993). Neonatal synesthesia: Implications for the processing of speech and faces. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 109–124). Dordrecht, The Netherlands: Kluwer Academic.
- Maurer, D., & Young, R. (1983). Newborns’ following of natural and distorted arrangements of facial features. *Infant Behavior and Development*, 6, 127–131.
- Mayes, L. C., & Carter, A. S. (1990). Emerging social regulatory capacities as seen in the still-face situation. *Child Development*, 61, 754–763.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90, 215–238.
- McGurk, H., & MacDonald, J. (1976, December 23/30). Hearing lips and seeing voices. *Nature*, 264, 746–748.
- Meltzoff, A. N., & Borton, R. W. (1979, November 22). Intermodal matching by human neonates. *Nature*, 282, 403–404.
- Meltzoff, A., & Kuhl, P. (1994). Face and speech: Intermodal processing of biologically relevant signals in infants and adults. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception* (pp. 335–369). Hillsdale, NJ: Erlbaum.
- Meltzoff, A. N., & Moore, M. K. (1993). Why faces are special to infants—On connecting the attraction of faces and infants’ ability for imitation and cross-modal processing. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 211–225). Dordrecht, The Netherlands: Kluwer Academic.
- Mendelson, M. J., & Ferland, M. B. (1982). Auditory–visual transfer in four-month-old infants. *Child Development*, 53, 1022–1027.
- Milewski, A. E. (1976). Infants’ discrimination of internal and external pattern elements. *Journal of Experimental Child Psychology*, 22, 229–246.
- Mills, A. E. (1987). The development of phonology in the blind child. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 145–161). Hillsdale, NJ: Erlbaum.
- Montague, D. (1995, March). *Infants’ discrimination of emotional expressions in a peekaboo interaction*. Paper presented at the meeting of the Society for Research in Child Development, Indianapolis, IN.
- Moon, C., Cooper, R., & Fifer, W. (1993). Infants prefer native language. *Infant Behavior and Development*, 16, 495–500.
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, 98, 164–181.
- Muir, D. W., Clifton, R., & Clarkon, M. (1989). The development of a human auditory localization response: A U-shaped function. *Canadian Journal of Psychology*, 43, 199–216.
- Muir, D. W., & Hains, S. M. J. (1993). Infant sensitivity to perturbations in adult facial, vocal, tactile, and contingent stimulation during face-to-face interactions. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 171–185). Dordrecht, The Netherlands: Kluwer Academic.
- Muir, D. W., Hains, S. M. J., & Symons, L. A. (1994). Baby and me: Infants need minds to read! *Cahiers de Psychologie Cognitive*, 13, 669–682.
- Muir, D. W., Humphrey, D. E., & Humphrey, G. W. (1994). Pattern and space perception in young infants. *Spatial Vision*, 8, 141–165.
- Neisser, U. (1994). Multiple systems: A new approach to cognitive theory. *European Journal of Cognitive Psychology*, 6, 225–241.
- Nelson, C. A. (1985). The perception and recognition of facial expressions in infancy. In T. Field & N. Fox (Eds.), *Social perception in infancy* (pp. 101–125). Norwood, NJ: Ablex.
- Nelson, C. A. (1987). The recognition of facial expressions in the first

- two years of life: Mechanisms of development. *Child Development*, 56, 58–61.
- Nelson, C. A., & Dolgin, K. (1985). The generalized discrimination of facial expressions by seven-month-old infants. *Child Development*, 56, 58–61.
- Nelson, C. A., & Horowitz, F. D. (1983). The perception of facial expressions and stimulus motion by 2- and 5-month-old infants using holographic stimuli. *Child Development*, 54, 868–877.
- Nelson, C. A., Morse, D. A., & Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child Development*, 50, 1239–1242.
- Olson, G. M. (1981). The recognition of specific persons. In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 37–59). Hillsdale, NJ: Erlbaum.
- Oster, H. (1981). "Recognition" of emotional expression in infancy. In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 85–125). Hillsdale, NJ: Erlbaum.
- Papousek, M., Bornstein, M. H., Nuzzo, C., Papousek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, 13, 539–545.
- Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and extension. *Infant Behavior and Development*, 18, 79–85.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.
- Premack, D., & Woodruff, G. (1978). Does a chimpanzee have a "theory of mind"? *Behavioral and Brain Sciences*, 1, 515–526.
- Proffitt, D. R., & Bertenthal, B. I. (1990). Converging operations revisited: Assessing what infants perceive using discrimination measures. *Perception and Psychophysics*, 47, 1–11.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rose, S. A., & Ruff, H. A. (1987). Cross-modal abilities in human infants. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 318–362). New York: Wiley.
- Rosen, W. D., Adamson, L. B., & Bakeman, R. (1992). An experimental investigation of infant social referencing: Mother's messages and gender differences. *Developmental Psychology*, 28, 1172–1178.
- Sackett, G. P. (1966, December 16). Monkeys reared in isolation with pictures as visual input: Evidence for an innate releasing mechanism. *Science*, 154, 1468–1473.
- Sackett, G. P. (1979). The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 623–649). New York: Wiley.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143–165.
- Scherer, K. R., Mounoud, P., Stern, D., Kappas, A., Zinetti, A., & Cesschi, G. (1994, June). *Emotional reactions to experimentally manipulated voice changes: A lab report on a cross-sectional study of infants between 5 and 14 months*. Paper presented at the International Conference on Infant Studies, Paris, France.
- Schiff, W. (1980). *Perception: An applied approach*. Boston: Houghton Mifflin.
- Sekuler, R., & Blake, R. (1994). *Perception* (3rd ed.). New York: McGraw-Hill.
- Sergent, J. (1987). Information processing and laterality effects for object and face perception. In G. W. Humphreys & M. J. Riddoch (Eds.), *Visual object processing: A cognitive neuropsychological approach* (pp. 145–174). Hillsdale, NJ: Erlbaum.
- Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in four- to six-month-old infants. *Developmental Psychobiology*, 25, 411–425.
- Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behavior and Development*, 18, 477–482.
- Sherrod, L. R. (1979). Social cognition in infants: Attention to the human face. *Infant Behavior and Development*, 2, 279–294.
- Slater, A., Earle, D. C., Morison, V., & Rose, D. (1985). Pattern preferences at birth and their interaction with habituation-induced novelty preferences. *Journal of Experimental Child Psychology*, 39, 37–54.
- Slater, A. M., Mattock, A., & Brown, E. (1990). Size constancy at birth: Newborn infants' responses to retinal and real size. *Journal of Experimental Child Psychology*, 49, 314–322.
- Soken, N. H., & Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Development*, 63, 787–795.
- Soken, N., Pick, A., Bigbee, M., Melendez, P., & Hansen, A. (1992, May). *Infant facial expression categories: Distinct categories of expression?* Paper presented at the International Conference on Infant Studies, Miami, FL.
- 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.
- Sparks, D. W., Ardell, L. A., Bourgeois, M., Wiedmer, B., & Kuhl, P. K. (1979). Investigating the MESA (Multipoint Electrotactile Speech Aid): The transmission of connected discourse. *Journal of the Acoustical Society of America*, 65, 810–815.
- Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, 8, 553–560.
- Spelke, E. S., & Cortelou, A. (1981). Perceptual aspects of social knowing: Looking and listening in infancy. In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition* (pp. 61–84). Hillsdale, NJ: Erlbaum.
- Spelke, E. S., & Owsley, C. J. (1979). Intermodal exploration and knowledge in infancy. *Infant Behavior and Development*, 2, 13–27.
- Spitz, R., & Wolf, K. (1946). The smiling response: A contribution to the ontogenesis of social relations. *Genetic Psychology Monographs*, 34, 57–125.
- Stack, D. M., & Muir, D. W. (1990). Tactile stimulation as a component of social interchange: New interpretations for the still-face effect. *British Journal of Developmental Psychology*, 8, 131–145.
- Stack, D. M., & Muir, D. W. (1992). Adult tactile stimulation during face-to-face interactions modulates five-month-olds' affect and attention. *Child Development*, 63, 1509–1525.
- Stein, B. E., Meredith, M. A., & Wallace, M. T. (1994). Development and neural basis of multisensory integration. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception* (pp. 81–106). Hillsdale, NJ: Erlbaum.
- Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 187–213). New York: Wiley.
- Stern, D. (1985). *The interpersonal world of the infant*. New York: Basic Books.
- Summy, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212–215.
- Summerfield, A. Q. (1979). Use of visual information in phonetic perception. *Phonetica*, 36, 314–331.
- Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3–51). London: Erlbaum.
- Thelen, E. (1986). Development of coordinated movement: Implications for early human development. In H. T. A. Whiting & M. G. Wade (Eds.), *Motor development in young children: Aspects of coordination and control* (pp. 107–124). Dordrecht, The Netherlands: Martinus Nijhoff.

- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47*, 466–472.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Eds.), *Before speech: The beginning of interpersonal communication*. Cambridge, England: Cambridge University Press.
- Tronick, E. Z., & Gianino, A. F. (1987). The transmission of maternal disturbance to the infant. In E. Z. Tronick & T. Field (Eds.), *Maternal depression and infant disturbance* (pp. 5–11). San Francisco: Jossey-Bass.
- Turkewitz, G. (1993). The influence of timing on the nature of cognition. In G. Turkewitz & D. A. Devenny (Eds.), *Developmental time and timing* (pp. 125–142). Hillsdale, NJ: Erlbaum.
- Turkewitz, G., Birch, H. G., & Cooper, K. K. (1972). Responsiveness to simple and complex auditory stimuli in the human newborn. *Developmental Psychobiology, 5*, 7–19.
- Turkewitz, G., & Devenny, D. A. (Eds.). (1993). *Developmental time and timing*. Hillsdale, NJ: Erlbaum.
- van Hooff, J. A. R. A. M. (1973). A structural analysis of the social behavior of a semi-captive group of chimpanzees. In M. von Cranach & I. Vine (Eds.), *Social communication and movement* (pp. 75–162). New York: Academic Press.
- Walden, T. A., & Baxter, A. (1989). The effects of context and age on social referencing. *Child Development, 60*, 1511–1518.
- Walden, T., & Ogan, T. (1988). The development of social referencing. *Child Development, 59*, 1230–1240.
- Walker, A. S. (1981). *Infants' intermodal perception of expressive behaviors*. Doctoral dissertation, Cornell University, Department of Psychology, Ithaca, NY.
- Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology, 33*, 514–535.
- Walker-Andrews, A. S. (1985, March). *The recognition of expressive behaviors across persons by infants*. Paper presented at the meeting of the Society for Research in Child Development, Montreal, Quebec, Canada.
- Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology, 22*, 373–377.
- Walker-Andrews, A. S. (1988). Infants' perception of the affordances of expressive behaviors. In C. K. Rovee-Collier (Eds.), *Advances in infancy research* (pp. 173–221). Norwood, NJ: Ablex.
- Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology, 3*, 55–75.
- Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expression by young infants. *Infant Behavior and Development, 6*, 491–498.
- Walker-Andrews, A. S., & Lennon, E. M. (1985). Auditory–visual perception of changing distance by human infants. *Child Development, 56*, 544–548.
- Walker-Andrews, A. S., & Lennon, E. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. *Infant Behavior and Development, 14*, 131–142.
- Walton, G. E., Bower, N. J. A., & Bower, T. G. R. (1992). Recognition of familiar faces by newborns. *Infant Behavior and Development, 15*, 265–269.
- Watson, J. S., & Ramey, C. T. (1972). Reactions to response-contingent stimulation in early infancy. *Merrill Palmer Quarterly, 18*, 219–228.
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Development and Behavior, 17*, 321–331.
- Werker, J. S., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.
- Wertheimer, M. (1961, November 24). Psychomotor coordination of auditory and visual space at birth. *Science, 134*, 1692.
- Wilcox, B., & Clayton, F. (1968). Infant visual fixation on motion pictures of the human face. *Journal of Experimental Child Psychology, 6*, 22–32.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.
- Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1977). Infant discrimination of facial expressions. *Child Development, 49*, 555–562.
- Zarbatany, L., & Lamb, M. (1985). Social referencing as a function of information source: Mothers versus strangers. *Infant Behavior and Development, 8*, 25–33.

Received January 31, 1995

Revision received July 30, 1996

Accepted August 9, 1996 ■