INFANCY, 2(4), 469–481

Perceiving the Real World: Infants' Detection of and Memory for Social Information

Arlene S. Walker-Andrews

Department of Psychology
Rutgers, The State University of New Jersey

Lorraine E. Bahrick

Department of Psychology
Florida International University

Research on infants' perception and memory for social information are discussed with respect to the concept of ecological validity. We argue that the use of dynamic, multimodal displays in familiar contexts is critical to understanding infants' developing perception. We also discuss the importance of generalizing appropriately from experimental to real-world contexts and of using converging methods in the investigation of infants' capabilities.

From birth, an infant is plunged into a world of other human beings in which conversation, gestures, and faces are omnipresent during the infant's waking hours. Moreover, these harbingers of social information are dynamic, multimodal, and reciprocal. It is no wonder that infants' early perceptual preferences include the human face, the human voice, animate motion, and events and interactions with these important social beings. Therefore, in the study of perceptual development we argue that it is paramount to include investigations of infants' recognition and responses to information provided by other human beings and to include exemplars of that information in the most potent (ecologically valid) form. After a discussion of what is meant by the term *ecological validity* in this context, we summarize re-

Requests for reprints should be sent to Arlene S. Walker-Andrews, Department of Psychology, 53 Avenue E, Rutgers University, Piscataway, NJ 08854-8040. E-mail: arlenewa@rci.rutgers.edu

sults from two different programs of research to illustrate the importance of ecological validity in the investigation of infants' capabilities.

To study the development of perception of complicated affordances of the environment, use of a number of converging methods is required. Specifically, if researchers want to know what individuals understand about the social and emotional information provided by other people, they need to approach the question from a number of angles using a number of procedures. For the study of infant development, this is especially critical because infants' behavioral repertoire is limited and the measures are necessarily indirect, requiring greater inferences about the bases of infants' responses. Madole and Oakes (1999) recently made a similar point regarding inferences about infants' abilities, specifically to form categories. They remarked on the "apparent gap ... about which set of responses reflects the child's actual categorical representation" (p. 273) and on the difficulty in designing studies to assess infants' abilities. They called for examining the contexts in which infants do and do not respond to particular categorical distinctions.

We, too, argue for the importance of measuring infants' abilities across contexts and, indeed, suggest that such designs permit infants to demonstrate capacities that may not otherwise be observable. *Ecological validity* is the term we use to describe combining experimental control with real-world applicability by designing experiments that measure infants' responses to information embedded in a rich context.

ECOLOGICAL VALIDITY

Brunswik (1955) coined the term *ecological validity* when he asked how the stimulus becomes a cue to the distal object that elicits it. It has undergone some redefinition and transformation since Brunswik's original conception. Today, ecological validity is used in a variety of ways (see Lickliter & Bahrick, this issue; Schmuckler, this issue) and refers to at least four important dimensions: (a) the nature of the stimuli (they should be events that capture important aspects of the natural environment rather than punctate images or disembodied sounds); (b) the nature of the context (it should be rich, nested in a larger setting, and containing important features of the natural environment to which generalizations are made); (c) the nature of the task or behavior (the task should require meaningful responses in the infant's natural repertoire); and (d) the nature of generalizations to the real world (their scope should be appropriate to the nature of the stimuli, context, and behaviors tested). An ongoing challenge to researchers has been to exercise adequate scientific control over the stimuli and context while maintaining ecological validity. Gibson (1969) emphasized the importance of the preceding dimensions as well as the need for scientific control:

> Mere observation of behavior in a real world may result in interesting anecdotes, but seldom provides enough stimulus control. Ideally, therefore, a laboratory situation which simulates a typical situation in the infant's world should be used. The situation must be controllable so that the information can be analyzed and describable so that it can be replicated, but it should be natural. An example of simulation of an ecologically significant environment for which an adaptively relevant response exists is provided by the visual cliff....
>
> The cliff apparatus permits more than [a] ... rough assessment, however, since it is possible to vary systematically the optical array of stimulation provided, so as to study the stimulus control of this avoidance. (p. 319)

Moreover, if one rejects the notion that inadequate sense inputs must be converted into veridical perceptions but accepts instead the idea that perception is a unified functional and intentional activity, an ecological approach that involves investigation of context is crucial. If perception is the apprehension of separate fragments of information, it matters little if that information is presented piecemeal. If, instead, the perceiver actively seeks structure in patterns of dynamic, multimodal, and contextually rich events, such stimulus materials are required to determine what the perceiver is capable of doing. In essence, if the whole is greater than the sum of its parts, then researchers are unlikely to be successful in understanding perception of the whole by attempting to examine the parts in isolation and then merely adding them together again. An illustration of this point can be drawn from a recent article by Partan and Marler (1999). They described how aromatic pyrazines and red and yellow coloration together produce aversion in chicks. However, one in the absence of the other does not lead to avoidance. The bimodal stimulation evokes a response that is not elicited by the unimodal compound, an example of an emergent property.

To enhance ecological validity in experimental studies of infants' perception, we argue that at least three aspects that typically characterize perception in the real world must be preserved: The stimulus objects or events provided for an observer should represent dynamic aspects of the world, they should be multimodal, and they should be placed in a meaningful context. In addition, one should use converging methods and measures to discern what infants perceive, given that the methods themselves create a sort of context that will vary. Finally, if researchers are to generalize perception and learning about real-world events, they should examine those processes as they are manifested in exploring and obtaining knowledge about that world.

RECOGNITION OF EMOTIONAL EXPRESSIONS

Research by one of us focused on infants' perception of expressive behaviors has been designed with ecological validity as a goal. In the Walker-Andrews laboratory, infants are presented with facial and vocal expressions to investigate their understanding of these expressions as communicating emotional information. Typically, expressions are presented on film to maintain stimulus control. To present posed, live expressions makes stimulus control a greater problem. Therefore, in a recent experiment (Montague & Walker-Andrews, 1996) the actress looked into a small mirror placed just above the infant's head to avoid emotion contagion and talked about emotion-producing events to induce her own emotions. None of these manipulations is perfect: The stimulus materials do not preserve all the properties of true emotional expressions. Nonetheless, after determining whether infants can discriminate the presented expressions or make intermodal matches or generalize across exemplars, aspects of the presentation are varied. The pitch, loudness, or timbre of the voice may be modified to determine whether these properties play a role in discrimination of vocal expressions. Alternatively, the actress may freeze her expression or, in the case of filmed exemplars, the face may be inverted or partially hidden from view. Synchrony may be disrupted so that infants cannot use temporal synchrony alone to match faces and voices in an intermodal preference task. The stimulus array is varied systematically and effects of these modifications are examined, but the initial stimulus is always the whole, elaborated (multimodal, dynamic) stimulus. Only later are infants given isolated pieces, rather than conducting separate experiments on infants' responses to each piece. One can then look at how the pieces fit back into the whole. Experiments that vary contextual information that adds meaning to the task of early recognition of emotion (cf. Baddeley, 1982) are underway.

In contrast, most research on the development of emotion perception has focused primarily on infants' ability to discriminate emotional expressions, with particular emphasis on charting the developmental course of discrimination of static facial expressions. Results indicate that from about 3 months of age infants can detect differences between static facial expressions of happiness, anger, fear, surprise, anger, and disgust (for reviews see Nelson, 1987; Walker-Andrews, 1997), although there is no indication that infants detect meaning in these expressions. Between 3 and 5 months infants discriminate among different vocal emotional expressions but only when these are concurrently presented with a facial expression (Caron, Caron, & MacLean, 1988; Walker-Andrews & Grolnick, 1983; Walker-Andrews & Lennon, 1991). Although these results are informative with respect to infants' ability to distinguish among different exemplars, whether infants extract any meaningful information from the expressions they appear to distinguish is unsettled. That is, although infants can discriminate facial expressions or vocal expressions, do these expressions convey meaning to the infants or are infants simply responding to differences in the set of the mouth or width of the eyes for facial expressions, or in the frequency (pitch) or timbre of a voice for vocal expressions?

Walker-Andrews and colleagues have tried to design experiments that allow us to come closer to providing infants opportunities to derive the meaning from expressive displays by using multimodal, dynamic, and more contextually rich materials. To return to one of the examples given previously, in an early experiment Walker-Andrews & Grolnick (1983) found that infants at 5 months could discriminate ongoing vocal expressions of emotion (happy and sad). The data from 3-month-olds were equivocal: Infants' responses were much more variable, and they seemed only to discriminate happy and sad voices when these were provided in a single direction—from sad to happy. Therefore, in a follow-up experiment, Walker-Andrews and Lennon (1991) examined infants' discrimination of these and other vocal expressions (happy, sad, and angry) and looked at whether it was critical for the infants that the vocal expressions be presented in a context. That is, would infants discriminate vocal expressions in the raw, so to speak, as disembodied voices emanating from a loudspeaker? Walker-Andrews and Lennon paired vocal expressions with affectively matching or affectively mismatching facial expressions or a checkerboard. In this case, 5-month-olds discriminated vocal expressions that were accompanied with facial expressions (of any sort) but did not show evidence for discrimination of vocal expressions that were accompanied by a checkerboard visual stimulus. However, researchers know that much younger infants can discriminate their mothers' voice from a stranger's (e.g., DeCasper & Fifer, 1980; Mehler, Bertoncini, Barriere, & Jassik-Gershenfeld, 1978), make pitch discriminations (e.g., Clarkson, Clifton, & Perris, 1988; Culp & Boyd, 1975; Trehub, Endman, & Thorpe, 1990), and distinguish one phoneme from another (e.g., Eimas, Siqueland, Jusczyk, & Vigorito, 1971). It seems as if infants failed to discriminate vocal expressions presented alone because without the contextual support of a face, infants did not abstract the specific information that allowed for discrimination of emotion in this task. Parameters such as frequency varied both within and between exemplars and did not provide an unambiguous foundation for discrimination (see also Dickson, 1998).

More recently, McCrink and Walker-Andrews (2000) embarked on a study in which they tried to determine the role of specific acoustic parameters that may be involved in infants' discrimination of vocal expressions. Walker-Andrews and McCrink used an intermodal preference technique in which two facial expressions are accompanied by a single vocal expression. The first trial was used as a baseline: How long will infants look at the affectively appropriate facial expression given a particular vocal expression? On the second trial, the vocal expression was modified (in this case its average fundamental frequency) to determine whether that information is critical to the infants' prior intermodal matching performance. The strategy was, again, to establish infants' abilities using rich, dynamic, multimodal stimulus materials in a controlled situation and to follow the original findings with probes to determine how such information contributes to the infants' abilities.

A second study on infants' perception of emotional expressions exploits familiarity as a context. Kahana-Kalman and Walker-Andrews (2001) examined whether contextual information such as person familiarity plays a role in the ability of young infants to perceive emotional expressions. The study compared the ability of 3.5-month-old infants to detect the correspondence between vocal and facial expressions when portrayed by their own mothers versus an unfamiliar woman.

Infants were presented two filmed facial expressions (happy and sad) accompanied by a single vocal expression. One group of infants observed facial and vocal expressions modeled by their own mothers (in synchrony), one group observed their mothers' facial and vocal expressions out of synchrony, and another group observed facial and vocal expressions portrayed by a stranger (in synchrony). Kahana-Kalman and Walker-Andrews (2001) recorded infants' looking time to determine whether they could make the intermodal match. In addition, they coded infants' emotional responses to the films to provide converging evidence about the meaningfulness of the stimulus expressions for the infants.

In brief, the infants in Kahana-Kalman and Walker-Andrews (2001) looked differentially to the appropriate, affectively matching facial expressions when the face and voice were of their own mothers, including infants who were presented synchronous and asynchronous pairings. Young infants apparently perceived the components of the happy and sad expressions of their own mothers as part of a unified, multimodal expression. Infants also looked preferentially to the happy facial expression when happy and sad films were presented simultaneously and in synchrony with the vocalization, indicating that infants found this filmed emotion the more compelling of the two films. In contrast, infants did not show by their looking patterns that they could detect the correspondence between vocal and facial displays of the same two emotions when these were portrayed by an unfamiliar woman. In addition, there were differences in infants' responsiveness to the filmed facial and vocal expressions. Global affective measures showed that when happy was the sound-specified emotion, infants were more expressive, showed greater variability of affective expression, and increased the number of alternating expressions. They were also rated as experiencing more positive affect and as more interested and engaged, particularly when the emotion displays were portrayed by their own mothers. Furthermore, more specific measures of infants' smiles and distress bouts showed that the infants who viewed their own mothers spent more time smiling at the films. Infants produced more full and bright smiles when happy was the sound-specified emotion and particularly when they viewed the happy expressions of their own mothers. The groups did not differ in duration of smiling when sad was the sound-specified emotion. The average duration of distress was significantly longer for infants who observed the unfamiliar woman.

To reiterate, infants showed intermodal matching of their mothers' facial and vocal expressions a full 4 months earlier than in prior studies that used an unfamiliar female as the model. Thus, familiarity with a particular stimulus or stimulus domain led to an earlier, enhanced responsiveness to the information available. Infants detected and responded to the affective correspondences in their own mothers' facial and vocal expressions even when synchrony relations between the

face and voice were disrupted, although they did not demonstrate a comparable ability for matching the synchronous expressions of an unfamiliar woman. Further, infants responded to the facial and vocal expressions with appropriate facial expressions of their own. For example, as they watched concurrent facial expressions and heard a single voice that matched one of them, they showed more smiles when the happy voice was played, and they showed more distress when an unfamiliar woman portrayed the emotions. Such patterns in the results are consistent with the view that one basis for infants' ability to detect correspondences between facial and vocal affective displays may be their ability to extract a common meaning. The results are also consistent with the general principles that familiarity leads to increased opportunity for differentiation and that differentiation proceeds from global properties to nested, more specific properties (see Bahrick, 2001; Gibson, 1969; Walker-Andrews, 1997), resulting in detection of more specific aspects of stimulation in more familiar stimuli. Overall, the findings demonstrate an early sensitivity to affective expressions that are dynamic, multimodal, and contextually relevant. Montague (2000) recently replicated and extended these results in an experiment that investigated young infants' perception of both their mothers' and their fathers' emotional expressions.

In summary, Walker-Andrews and colleagues have examined infants' abilities to discriminate and recognize emotional expressions by testing infants with a number of methods (recording looking time and affective responsiveness in visual habituation, intermodal preference, and mother–infant interaction studies) and by designing the experiments to be maximally ecologically valid. Infants view emotional expressions that are presented as multimodal, dynamic, contextually rich events, rather than unimodal, static, and detached. Infants have demonstrated that they recognize the expressions portrayed by familiar persons (their mothers) earlier than those of an unfamiliar person and that they discriminate vocal expressions only when these are embedded in a relevant context (the face). Thus, when ecological validity is enhanced by using dynamic, multimodal events that are familiar, infants show an ability to abstract the invariant information that potentially carries the meaning of an emotional expression.

MEMORY FOR FACES AND FACES IN ACTION

Studies of infants' perception and memory for faces also underscore the importance of using dynamic, multimodal presentations (Bahrick, Gogate, & Ruiz, 2001) and highlight the disparity between results obtained from dynamic versus static presentations. Infant memory for videos of women performing simple actions was investigated. Five-month-old infants were familiarized with silent video films showing close-ups of women's faces and shoulder areas as they performed different actions such as brushing their hair, blowing bubbles, or brushing their teeth. The women were chosen so as to have discriminably different faces: One was White with very long, light brown hair; another was from India with shoulder length wavy hair; and the third was Asian with long, straight black hair. Infants were familiarized with one woman performing one action. One minute later, and again 1 to 2 months later, the infants were given novelty preference tests to assess memory for the action and the face. Results indicate that after the 1-min and the 1- to 2-month delays, infants remembered the actions. They preferred to look at the novel action after the 1-min delay and the familiar action after the 1- to 2-month delay. This shift in preference was consistent with prior research from Bahrick's laboratory (Bahrick, Hernandez-Reif, & Pickens, 1997; Bahrick & Pickens, 1995; see also Courage & Howe, 1998; Spence, 1996). In contrast, infants showed no evidence of memory for the faces in the overall results at either delay.

Given that infants of this age have demonstrated discrimination and short-term memory for faces under a variety of conditions (still, photos, actual faces, bimodal speaking) in prior studies, it was hypothesized that the simple actions were so salient that infants failed to attend to the appearance of the faces. That is, the actions became "foreground" whereas the distinctive features specifying the identity of the faces became "background" (see Bahrick & Lickliter, 2000, for an analogous discussion of the roles of unimodal and multimodal information). To address this issue, Bahrick et al. (2001) conducted a control study to determine if infants could discriminate among the faces when they were static. The same faces were presented to infants, this time showing several static views of each woman performing the action (taken from the videos of the moving displays). Results indicated significant discrimination of the faces after the 1-min delay but not the 1- to 2-month delay.

These findings highlight the importance of dynamic information and the disparity between dynamic and static information for guiding attention, perception, and memory in early infancy. When presented with persons engaged in repetitive activities, infants focus on the actions to a greater extent than the features of the face. When there is no action, the static features of the face become more salient and are discriminated. Whether this pattern changes developmentally is currently under investigation. These findings demonstrate that results of studies on face perception are markedly different as a function of whether the faces are still or moving. Further, they suggest that to generalize to attention, perception, or memory for faces in the real world, researchers must present faces as they occur, in the context of dynamic, moving, or speaking individuals. To generalize to the perception of static faces, however, one must present nonmoving faces as stimuli.

Further evidence for the importance of dynamic information comes from a study of speech perception by Gogate and Bahrick (1998). In this study 7-month-old infants were taught a relation between two verbal labels and two distinctive looking objects. Infants were only able to learn the relation when the objects were moved in synchrony with the speech sounds, not when they were static or were moved out of synchrony with the sounds. A companion study (Gogate, Bahrick, & Watson, 2000) found that mothers use a great deal of synchrony between verbal labels and object motions when they teach their infants new names for objects. These findings demonstrate the importance of dynamic information in directing attention to important relations and facilitating learning of verbal labels for young infants.

CONCLUSIONS

Were it merely an empirical question—Is it better to design experiments in such a way that they are ecologically valid?—the evidence would be compelling. Infants in the research summarized here responded differentially to information when it was dynamic versus static, multimodal versus unimodal, familiar versus unfamiliar (see also Lickliter & Bahrick, this issue). These dimensions apparently have an important effect on memory, perception, and attention in infants. For example, there were clear benefits for perception and memory of dynamic social information generated from dynamic, multimodal, and contextually rich presentations. In contrast, repetitive activities are apparently sufficiently salient to infants that perception and memory for the configuration or appearance of these dynamic stimuli (e.g., facial identity) is diminished. Researchers drawing generalizations from experimental settings to the real world should thus take these dimensions into account. For example, if one wishes to generalize to dynamic events, the stimuli must be dynamic. However, if one wishes to explore perception of static dimensions—for example, the appearance of faces—and to generalize to identification of static faces, the stimuli must be static. In other words, to be ecologically valid, generalizations must be grounded in the important dimensions of the stimuli and context that "make a difference."

It is also clear that the critical issues are not just of generalizability or sensitivity but of definition. If one believes that perception results from interpreting sensory processes, as the mechanistic theories prevalent in psychology presume, then perception must depend on sensation. These theories of perception require that the observer detect the stimuli, interpret the stimuli to mean something, and then that he or she use the interpretation to change behavior. In this view, sensitivity to and awareness of punctate stimulation is key. Experiments in which infants are exposed to the isolated pieces of a stimulus event are thought, therefore, to yield findings that will allow the experimenter to describe how infants respond to the whole event—a combination of the features. If one assumes instead that perception results from an active process of detecting meaningful patterns of stimulation, then perception does not depend on sensation in the same way. From this ecological perspective, it is not useful to describe how the infant makes sense of the whole based on perception of to-be-aggregated features. Rather, this approach asks whether an observer responds to objects and events in their totality, as dynamic, multimodal, contextually embedded objects that are meaningful. The approach of presenting the whole (dynamic, multimodal, context-rich) object or event becomes optimal. Then one can progress to detaching various parts to determine whether differentiated behavior disappears.

This strategy can be found in the studies outlined here, as, for example, when infants were asked to discriminate vocal expressions that were presented along with facial expressions (Walker-Andrews & Grolnick, 1983; Walker-Andrews & Lennon, 1991). Five-month-olds failed to show evidence for discrimination when a vocal expression was accompanied by a checkerboard stimulus but showed discrimination when the voice was accompanied by a facial expression, whether or not that expression was consonant (same emotion). In our view, the presence of a facial expression provided a social context that contributed to infants' detection of emotion information in the vocal expressions.

Dynamic systems approaches also assume that behavior occurs in complex, variable environments and requires the capacity to use experience in functional ways. The information available to perception systems is well described in dynamic system terms (Butterworth, 1993) given the focus on the transitions among components and adaptations of a system. Behavior and perceptual awareness are the ways organisms regulate their encounters with the environment, leading to the question of how organisms coordinate their behavior in terms of the affordances they detect. Researchers can only benefit from such an enterprise—creating new descriptions of stimulus information and examining responsivity to these.

Given the data summarized here, we conclude that young infants require dynamic, multimodal displays to best demonstrate their perceptual competencies for social events. What young infants first perceive in such dynamic displays are superordinate aspects and patterns of stimulation over time. For example, faces, voices, emotional expressions, and patterns of skeletal articulation are perceived in the world to arise from a unified, embodied source. Results from studies of infants' recognition of emotional expressions and their perception and memory for faces, actions, and labels indicate that infants solve the perceptual problems researchers present to them at the level researchers present them. That is, one can demonstrate infants' sensitivity to disembodied acoustic parameters or static visual patterns using simplistic stimuli. In this case, however, when infants show sensitivity to these dimensions we suggest it provides us with an assessment of their ability to perceive these dimensions but not an assessment of their ability to perceive embodied acoustic parameters or dynamic visual patterns. In contrast, one can examine their ability to use such information as it typically occurs (as one facet of a vocal expression or environmental event). In this case scientists come much closer to finding evidence for how infants come to understand, interact with, and adjust their attention to the real world of objects, events, and social beings.

In closing, developmental researchers have learned much about infants' perceptual sensitivity and their discrimination of visual patterns, faces, sounds, and so forth. Traditionally, researchers have examined infants' responsiveness to unimodal visual or acoustic information. This separation by modes is partially historical, stemming from theories of perception that described the process by considering discrete stimuli setting off receptors giving rise to individual sensations that were integrated in some way, such as by association. If instead research is guided by a view of the infant as an active perceiver of meaningful patterns of stimulation that occur over time and in three-dimensional space, researchers must investigate perception as it is embedded in a rich, multimodal context. In this research, we find that infants show more representative capabilities when information is dynamic and ordered, events are multimodal, and objects are set in contexts. This approach allows scientists to learn more about how perception functions and develops in the real world.

ACKNOWLEDGMENTS

A preliminary version of this article was presented at the meetings of the Society for Research in Child Development, April 1999, Albuquerque, NM.

The research was supported by grants awarded to Arlene S. Walker-Andrews (NSF SBR-9408993) and Lorraine E. Bahrick (NICHD R01 HD25669).

We wish to thank Linnea R. Dickson, Ronit Kahana-Kalman, Lakshmi Gogate, and Ivonne Ruiz for their assistance on the research. We also thank the infants and their parents who participated.

REFERENCES

Baddeley, A. (1982). Domains of recollection. *Psychological Review, 58*, 708–729.

Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology, 79*, 253–270.

Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2001). *Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events.* Manuscript submitted for publication.

Bahrick, L. E., Hernandez-Reif, M., & Pickens, J. N. (1997). The effect of retrieval cues on visual preferences and memory in infancy: Evidence for a four-phase attention function. *Journal of Experimental Child Psychology, 67*, 1–20.

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

Bahrick, L. E., & Pickens, J. N. (1995). Infant memory for object motion across a period of three months: Implications for a four-phase attention function. *Journal of Experimental Child Psychology, 59*, 343–371.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217.

Butterworth, G. (1993). Dynamic approaches to infant perception and action: Old and new theories about the origins of knowledge. In L. B. Smith & E. Thelen (Eds.), *A dynamic systems approach to development* (pp. 171–188). Cambridge, MA: MIT Press.

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.

Clarkson, M. G., Clifton, R. K., & Perris, E. E. (1988). Infant timbre perception: Discrimination of spectral envelopes. *Perception and Psychophysics, 43,* 15–20.

Courage, M. L., & Howe, M. L. (1998). The ebb and flow of infant attentional preferences: Evidence for long-term recognition memory in 3-month-olds. *Journal of Experimental Child Psychology, 10,* 26–53.

Culp, R. E., & Boyd, E. F. (1975). Visual fixation and the effect of voice quality and content differences in 2-month-old infants. In F. D. Horowitz (Ed.), 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), 78–91.

DeCasper, A. J., & Fifer, W. P. (1980, June 6). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174–1176.

Dickson, L. R. (1998). *Infants' perception of facial and vocal expressions of affect: The significance of multimodality, intermodal discrepancies, and contexts.* Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971, January 22). Speech perception in infants. *Science, 171,* 303–306.

Gibson, E. J. (1969). *Perceptual learning and development.* New York: Appleton-Century-Crofts.

Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology, 69,* 133–149.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 71,* 878–894.

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

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review, 19,* 263–296.

McCrink, K. C., & Walker-Andrews, A. S. (2000, July). *The role of pitch in infants' intermodal matching of facial and vocal expressions.* Poster session presented at the 12th Biennial International Conference on Infant Studies, Brighton, England.

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gershenfeld, D. (1978). Infant recognition of mother's voice. *Perception, 7,* 491–497.

Montague, D. P. F. (2000, July). *Fathers, mothers, and infants: familiarity and early perception of emotional communication.* Poster session presented at the 12th Biennial International Conference on Infant Studies, Brighton, England.

Montague, D. P. F., & Walker-Andrews, A. S. (1996, April). *Infants' expectations and emotional responsivity during peekaboo interactions.* Paper presented at the 10th Biennial International Conference on Infant Studies, Providence, RI.

Nelson, C. A. (1987). The recognition of facial expressions in the first two years of life: Mechanisms of development. *Child Development, 56,* 58–61.

Partan, S., & Marler, P. (1999, February 26). Communication goes multimodal. *Science, 283,* 1272–1273.

Spence, M. (1996). Young infants' long-term auditory memory: Evidence for changes in preference as a function of delay. *Developmental Psychobiology, 29,* 685–695.

Trehub, S. E., Endman, M. W., & Thorpe, L. A. (1990). Infants' perception of timbre: Classification of complex tones by spectral structure. *Journal of Experimental Child Psychology, 49,* 300–313.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological Bulletin, 121,* 437–456.

Walker-Andrews, A. S., & Grolnick, W. (1983). Infants' discrimination of vocal expressions. *Infant Behavior and Development, 6,* 491–498.

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.