Journal of Experimental Child Psychology **79,** 253–270 (2001)
doi:10.1006/jecp.2000.2588, available online at http://www.idealibrary.com on IDEAL®

Increasing Specificity in Perceptual Development: Infants' Detection of Nested Levels of Multimodal Stimulation

Lorraine E. Bahrick

Florida International University

This research assessed the development of infants' sensitivity to two nested amodal temporal relations in audible and visible events. Their detection of global temporal synchrony between visible and audible impacts and internal temporal structure nested within each impact specifying object composition (single versus compound objects) was assessed. Infants of 4, 7, and 11 weeks of age were habituated to a single and a compound object striking a surface and then received test trials depicting a change in synchrony or object composition. Results indicated an interaction between age and condition where sensitivity to synchrony was present by 4 weeks and remained stable across age, whereas sensitivity to composition emerged later, by 7 weeks, and increased dramatically with age. These findings converge with other recent findings to illustrate a pattern of increasing specificity in the development of perception, where infants first detect global and later detect embedded relations. The early sensitivity to global relations may provide an organizational framework for development by focusing infant attention on unitary events, guiding and constraining further exploration, and buffering infants from learning incongruent relations.

Key Words: infant perception; intersensory perception; intermodal perception; increasing specificity; auditory–visual perception; temporal synchrony; developmental trends.

Objects and events can be characterized as having hierarchically organized properties. Some properties are global and abstract, whereas others are more specific and nested within global properties. For example, we live in a world of objects and events that presents an array of information to all the senses concurrently. Some of this information is "amodal" and not tied to a particular sense modality. In this sense, it is global and abstract. For example, the temporal synchrony, rhythm, and tempo common to the sights and sounds of a ball bouncing are amodal. Other information is modality-specific and perceivable through only one sense modality. In this sense it is more specific. For example, the pitch of the

This research was supported by a National Institute of Child Health and Human Development grant (RO1 HD25669). Some of these data were presented at the International Conference on Infant Studies in Atlanta, GA, April 1998, and in Providence, RI, April 1996. Special thanks are extended to Ivonne Fernandez, Maria Hernandez-Reif, and Brenda Lundy for their help in participant testing and recruitment, and to Lakshmi Gogate for her help with data analyses.

Address correspondence and reprint requests to Lorraine E. Bahrick, Department of Psychology, Florida International University, Miami, FL 33199.

0022-0965/01 $35.00

impact sounds and the color and pattern of the ball are modality-specific. Gibson (1969) proposed a principle of increasing specificity in the domain of perceptual development. According to this view, development occurs in order of increasing specificity such that initially infants detect abstract, global properties of events and later they learn to differentiate finer and more detailed aspects of stimulation.

A great deal of research has now accrued demonstrating that infants are adept perceivers of global and nested properties of events. However, little research has explored the developmental trajectory of infants' sensitivity to different levels of stimulation. For example, research has established that infants are skilled perceivers of amodal properties, including temporal and spatial aspects of events, even in the first half-year of life (see Lewkowicz & Lickliter, 1994, for a review). They can perceive the relation between a face and a voice on the basis of the common emotional expression (Walker, 1982; Walker-Andrews, 1997). Infants can relate moving objects to their impact sounds on the basis of temporal synchrony (Bahrick, 1983, 1987, 1988; Lewkowicz, 1992, 1996), their common tempo (Spelke, 1979), rhythm (Bahrick & Lickliter, 2000; Mendelson & Ferland, 1982), colocation (Fenwick & Morrongiello, 1998; Morrongiello, Fenwick, & Nutley, 1998), and changing distance (Pickens, 1994; Walker-Andrews & Lennon, 1985). They can also detect temporal information specifying object substance (Bahrick, 1983) and composition (Bahrick, 1987, 1988, 1992).

Some amodal properties can be characterized as more global, whereas others are nested and more specific. For example, temporal synchrony, the coincidence between the sights and sounds of an object's impacts, has been described as one of the most global, abstract temporal relations (Bahrick, 1992, 1994; Bahrick & Lickliter, 2000; Bahrick & Pickens, 1994). It can be detected only by abstracting information across different sense modalities by relating information about audible and visible changes over time. Synchrony can therefore be characterized as the most global amodal property because it is inherently relational and abstract. Other amodal properties such as tempo, rhythm, duration, and temporal information specifying object composition and substance are also global properties, but are considered to be more specific than synchrony because they can be perceived through a single sense modality (although they are not *specific* to any one sense modality). Further, if they are conveyed simultaneously in more than one modality, synchrony is necessary for accurate perception of these amodal properties.

It has been proposed that temporal synchrony is one of the first types of information that infants abstract (Bahrick & Pickens, 1994; Lewkowicz, 1989, 1996, 2000). This is adaptive because sights and sounds that are synchronous typically specify unitary events. Thus, infant sensitivity to this information can foster processing of audible and visible information that belongs together (rather than disjointed sights and sounds) and can then promote further differentiation of embedded properties in unitary, meaningful events. This may serve as an important basis for early veridical perceptual organization.

Infants are also excellent perceivers of modality-specific properties. For example, they can detect color (e.g, Bornstein, 1978; Dannemiller, 1989; Peeples &

Teller, 1975; Teller & Bornstein, 1987) and pattern in two-dimensional visual forms, static and dynamic faces, and facial expressions (e.g., Barrera & Maurer, 1981; Caron, Caron, & MacLean, 1988; Fagan, 1972, 1976; Fantz, Fagan, & Miranda, 1975; Nelson & Horowitz, 1983). They can also perceive pitch and/or timbre of voices (e.g., DeCasper & Fifer, 1980; DeCasper & Prescott, 1984; Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991) and tones (e.g., Clarkson & Clifton, 1984; Clarkson, Martin, & Miciek, 1996). Less is known, however, about infants' sensitivity to intersensory *relations* between modality-specific properties. Unlike amodal relations which can be directly perceived and remain invariant across contexts, modality-specific relations must be learned. They are arbitrary and vary across events and contexts. For example, the word "pen" may stand for a writing instrument in English but not Spanish; a blue ball may make a high-pitched impact sound hitting one surface, but a deeper sound hitting another; a particular person may typically speak with a clear voice, but at other times with a raspy voice. These relations can be described as the most specific level because they are modality-specific as well as context-specific (Bahrick, 1994). Infants can learn to relate modality-specific properties across the senses, but this appears to be somewhat more difficult and emerges later than relating amodal properties across the senses (Bahrick, 1992, 1994). For example, by 7 months, but not at 3 or 5 months, infants can relate the pitch of an impact sound with the color and shape of an object (Bahrick, 1994). Seven-month-olds can also relate the color of an object with its taste (Reardon & Bushnell, 1988), and 4-month-old female infants can relate the odor of an object with its appearance (Fernandez & Bahrick, 1994). Further, relating modality-specific properties across the senses appears to be facilitated by the presence of temporal synchrony and/or other amodal relations. For example, 7-month-old infants were able to relate the appearance of an object with a specific speech sound if the speech sounds and objects were presented synchronously, but not if they were presented asynchronously or without motion (Gogate & Bahrick, 1998). In one study, even neonates were able to relate a sound and an object when they were presented contingent on the infant's behavior but not when they were presented noncontingently (Slater, Quinn, Brown, & Hayes, 1999).

The research reviewed above on infants' detection of amodal and modality-specific audio–visual relations has yielded important insights into infant capabilities. However, it cannot elucidate the nature of the developmental trajectory or the basis for developmental change. Since the time Gibson initially proposed the principle of increasing specificity more than 30 years ago, few systematic attempts to empirically test this principle have been made (but see Lewkowicz, 1989, 2000; Walker-Andrews, 1997, for relevant reviews). Systematic research aimed at uncovering the bases of development and assessing changes across age in specific abilities is necessary.

To address the issue of increasing specificity directly requires a systematic body of research assessing developmental trends within specific domains using methods, tasks, and stimuli that are controlled across age. Further, a priori hypotheses regarding developmental patterns within a domain are necessary for explicitly evaluating the principle of increasing specificity. For example, by standardizing methods and procedures across age, and by using events that provide multiple levels of information, one can directly test at what age detection of different types of information available in a given event emerge. Given a priori definitions about which relations are more global and which are embedded, specific hypotheses can be systematically evaluated.

Bahrick (1992, 1994) recently conducted this type of research to assess the principle of increasing specificity in the domain of intersensory development. These studies assessed whether detection of amodal relations would developmentally precede the detection of arbitrary, modality-specific relations. Audio–visual events depicting objects striking a surface were chosen so as to convey intersensory relations of three levels of specificity. Amodal temporal synchrony between the sights and sounds of the objects' impacts constituted the most global level of temporal structure. A second more specific and embedded amodal relation was the internal temporal structure (microstructure) that characterizes the nature of each synchronous impact (Bahrick 1983, 1987, 1988). This temporal structure common to the type of sound and motion can convey information about an object's substance, composition, weight, size, and number. In this research, infants' detection of composition information was tested. A single object striking a surface produces a single, abrupt impact sound and has a single, abrupt visual trajectory change. In contrast, an object composed of many smaller elements (a "compound" object) produces a more complex sound with a more prolonged onset and has a correspondingly more complex and prolonged visual trajectory change. The internal temporal microstructure is common across the visual and auditory stimulation and specifies the object's composition. A third, more specific intersensory relation provided by the events was the modality-specific, arbitrary relation between the pitch of the sound and the color/shape of the object.

Three-month-old infants received alternating exposures to a single and a compound event with their natural sounds until they reached habituation. They were then tested under one of three conditions to assess detection of a change in amodal synchrony (from synchronous to asynchronous), amodal composition (from matched to mismatched sounds), or arbitrary pitch–color/shape relations. Results (Bahrick 1992) indicated that the 3-month-olds showed significant visual recovery to the change in synchrony and composition relations, but not to the change in the arbitrary, modality-specific relation between the pitch of the impact sound and the color/shape of the object. Thus, there appears to be a developmental lag between detection of amodal and arbitrary intersensory relations. This lag did not appear to depend on limitations to the infants' perceptual abilities. A control study demonstrated that the 3-month-olds were able to detect all the color/shape and pitch changes used. However, they did not relate them across the senses until 7 months of age (Bahrick, 1994). Converging evidence for this developmental lag between the detection of amodal and arbitrary intersensory relations comes from a set of studies on visual-tactile perception of objects (Hernandez-Reif & Bahrick, 2001). Infants of 4 and 6 months of age detected amodal information for object shape, but only the 6-month-olds were able to detect the arbitrary shape–color/pattern relation.

These findings of a developmental lag support the principle of increasing specificity. Further, this developmental lag is likely to be adaptive in promoting accurate perception of objects and events. By detecting amodal relations developmentally prior to arbitrary intersensory relations, infants can acquire knowledge about persistent, context-free intersensory relations prior to learning about arbitrary relations that vary from one context or event to another. This can serve to enhance appropriate generalizations and minimize idiosyncratic generalizations about the relations between sights and sounds.

Establishing increasing specificity as a general developmental principle requires converging evidence across a variety of domains, methods, and procedures. As a step in this direction, the present study presents another systematic effort at directly testing the principle of increasing specificity, this time in the domain of global and nested amodal relations. Using the single and compound events employed in the above studies, the present research assessed whether infants would detect global synchrony relations developmentally prior to nested intersensory information specifying object composition. Given that infants of 3 months detected both the synchrony and composition relations in our prior research, younger infants were tested in the present investigation to examine the emergence of these abilities. If synchrony is in fact a fundamental basis for linking sights and sounds, one should find that detection of synchrony developmentally precedes detection of other embedded relations. In this manner, sensitivity to synchrony could appropriately constrain further differentiation of events such that only multisensory stimulation that belonged together was subsequently processed together. Without such an organizing principle, early perception would likely be haphazard, with infants attending to some visual components of events while listening to unrelated auditory streams.

Method

Subjects. Ninety-six infants (49 males and 47 females) participated. There were 20 4-week-olds (13 males and 7 females) with a mean age of 28.3 days ($SD = 3.61$), 28 7-week-olds (12 males and 16 females), with a mean age of 50 days ($SD = 3.95$), and 48 11-week-olds (24 males and 24 females), with a mean age of 77 days ($SD = 5.64$). (The N is larger for the 11-week-old group because the data of 24 infants who participated in another study afterward was also included. It was decided to combine these data because the larger N gives a more accurate estimate of means and greater power for observing within group differences. Thus, the relative contributions of synchrony and composition information could be better assessed with the larger sample.)

In all, 48 additional subjects were tested but their data were not included, 24 because of excessive fussiness ($N = 3$, 3, and 18 at 4, 7, and 11 weeks, respectively), 4 because of equipment failure or experimenter error ($N = 3$, 0, and 1 at 4, 7, and 11 weeks, respectively), and 11 for falling asleep or excessive drowsiness during the procedure ($N = 2$, 6, 3 at 4, 7, and 11 weeks, respectively). Further, 1 4-week-old was rejected for failure to meet the attention criterion, and 8 infants (4 7-week-olds and 4 11-week-olds) were rejected for failure to habituate (see Procedures for details on attention and habituation criteria). Infants were recruited through local birth records. They were all apparently healthy with no complications during delivery and had Apgar scores of at least 9. The sample was primarily middle class. Parents were required to have at least 12 years of education to participate.

Stimulus materials. Color video films of two pairs of naturalistic audio–visual events were used (see Bahrick, 1992, 1994). One pair depicted a large, single, orange hexagonal nut versus a cluster of smaller, orange, hexagonal nuts, and the other pair depicted a large, single, yellow, circular washer versus a cluster of smaller, yellow, circular washers. They were chosen so as to portray two types of object composition. One member of each pair depicted a single, large metal object (single object) and the other a cluster of smaller, metal objects similar in appearance (compound object). Each object and object cluster was suspended from a string and filmed crashing against a surface in a erratic temporal pattern, producing natural impact sounds at a rate of approximately 40 impacts per minute. The impact sounds of both the single and compound objects were abrupt and discrete and differed primarily in terms of internal temporal structure. The single objects each produced a single, discrete impact sound and had a single, discrete visible trajectory change at impact. The onset of the sound and visual trajectory change was abrupt. The compound objects each produced a more complex impact sound with a more complex, prolonged sound onset and had a correspondingly complex visual trajectory change at impact. Thus, the sounds and sights of impact were united by both a temporal synchrony relation and a common temporal microstructure specifying the object's composition.

For some test trials (composition change trials), the wrong sounds were synchronized with the object motions; that is, the sounds of the single object were synchronized with the visual impacts of the compound object and vice versa (see Bahrick, 1992, 1994). These events were created by filming (with two cameras) both the single and compound objects simultaneously, one through the window of a sound-attenuated chamber, and the other through the window of a chamber that was not sound-attenuated. An experimenter, holding one object in each hand, abruptly and simultaneously struck each object against its respective surface, carefully moving both hands in synchrony with one another. In this manner, the soundtrack to one object could be precisely aligned with the motions of the other. The mean temporal disparity (calculated on the basis of 10 randomly selected impacts) between the onset of a sound and the point of visual contact between the object and surface was .009 s (range = 0–.03 s, or 0–1 frame). In contrast, some test trials depicted the correct sounds presented asynchronously (synchrony change trials). For these trials, the soundtrack to each event was dubbed onto the visual portion of the event, beginning in a random position, such that the sounds and visual impacts were unsystematically out of synchrony (given the erratic temporal pattern of the impacts). These asynchronous events had a mean temporal disparity of .8 s with a range of 2.84 to .06 s across a sample of 10 randomly chosen impacts.

Since these stimulus events had not previously been used with infants as young as 4 weeks, and no literature existed assessing whether infants of this age could detect the visual or the auditory temporal microstructure specifying object composition, a preliminary study was conducted to address this question. Eight infants ($M = 23.6$ days, $SD = 4.1$) were habituated to an audio–visual event ($N = 4$ single and $N = 4$ compound) and then received four test trials, each depicting one event. Two test trials were designed to assess discrimination of the visual change from a single to a compound object or vice versa, and two assessed discrimination of the auditory change from a single to a compound impact sound or vice versa. For the visual test, the soundtrack remained unchanged, but the sounds were synchronized with movements of an object of a new composition. For the auditory test, the object remained unchanged, but its movements were synchronized with a soundtrack of a new composition. Results indicated robust visual recovery to the visual [$t(7) = 5.47$, $p = .001$] and auditory changes [$t(7) = 3.71$, $p = .008$] specifying object composition. All eight infants showed positive visual recovery scores to both the auditory and visual changes. Thus, infants as young as 3–4 weeks were able to discriminate between the single and compound objects and the single and compound sounds used in this study.

Apparatus. Infants were seated in a standard infant seat approximately 55 cm from a 19-in. (Panasonic BT-S1900N) video monitor, surrounded by black posterboard. Two apertures in the posterboard, one to the upper right and the other to the upper left side of the monitor, allowed observes to view the infants' visual fixations.

The audio–visual events were videotaped using a Panasonic (WV3170) color video camera and a Sony (EMC-150T) remote microphone. They were edited and presented with a Panasonic (VHS NV-A500) edit controller that was connected to four Panasonic video decks (NV-8500s and AG-6300s). With four decks we were able to switch among the two habituation and two test displays without the noise or extra time resulting from changing cassettes. The soundtracks were presented from a speaker just beneath the video screen at a level of approximately 65 dbA measured in the area of the infant seat.

A trained observer, unaware of the infant's condition, monitored visual fixations by pressing a button while the infant fixated the video image. The button box was connected to a PC that was programmed to record visual fixations online, signal as soon as the infant had looked away 1.0 s to end a trial, signal if the infant reached the maximum looking time for a given trial of 45 s, and signal when the habituation criterion had been reached. The experimenter who controlled the video displays, received the signals through a headphone connected to a small speaker. A permanent record of the infant's visual fixations was created on the computer during the experiment. The observations of the primary observer controlled the audio–visual presentations, whereas those of a secondary observer (shielded from view of the primary observer) were recorded for later calculation of interobserver reliability.

Procedure

Infants were tested using an infant-controlled habituation procedure similar to that of Bahrick (1992, 1994) to determine whether they could detect the audio–visual relations of temporal synchrony or temporal microstructure specifying object composition. Half the infants in each age group were randomly assigned to the synchrony change condition and half to the composition change condition, with the exception of those in the 7-week-old group. In this age group, an additional four infants participated in the composition change condition. Half the subjects within each of these groups received the event pair depicting the yellow washers and the other half the orange nuts. Each infant was habituated to an alternating sequence of a single and a compound event comprising a pair. The events were presented in synchrony with their natural sounds. Then for test trials, infants received the same two events; however, the auditory and visual components were presented in a novel relation to one another. That is, for the synchrony change condition, the sounds were now presented out of synchrony with the visible motions of the objects. The films and soundtracks were misaligned, and since their temporal patterns were erratic, they were unsystematically out of phase with one another. For the composition change condition, the films were presented in synchrony with the objects of the wrong composition. That is, the impact sounds of the single object were synchronized with the motions of the compound object and vice versa. Thus, during the test trials, both the films and the soundtracks were familiar to the infants, only the relation between them was novel.

In general, the habituation procedure consisted of an initial control trial, four mandatory habituation trials, and was terminated after the infant reached the habituation criterion of a 50% decrement in visual fixation level on two consecutive trials relative to the infant's initial fixation level on the first two habituation trials. Then, two no-change posthabituation trials were presented. They served to establish a more conservative criterion for habituation by reducing chance habituation and taking into account spontaneous regression effects by assessing visual recovery in relation to these trials (see Bertenthal, Haith, & Campos, 1983, for a discussion of regression effects). Following this habituation sequence, infants received two test trials (one with each event) to assess visual recovery and then a final control trial. The initial and final control trials depicted a toy turtle with arms spinning, creating a whirring sound. Performance on these trials was compared against a criterion to assess fatigue or inability to show visual recovery (see Bahrick, 1992, 1994). The visual fixation on the final control trial was compared with that on the initial control trial. If the final level was less than 10% of the initial level, the infant's data were rejected ($N = 1$ at 4 weeks). The remaining infants showed a large visual recovery on the final control trial (median = 70%

at 4 weeks, 99% at 7 weeks, and 102% at 11 weeks of age) relative to the initial fixation level. A second criterion was also adopted (see Bahrick, 1992, 1994) to ensure that infants had, in fact, habituated to the audio–visual events. If the infant's mean posthabituation fixation level exceeded that of their mean initial fixation level (baseline), the data were rejected from the study ($N = 0$, 4, and 4 at 4, 7, and 11 months respectively).

The data of infants in this study were more difficult to collect than those of prior studies with older infants (Bahrick, 1992, 1994), and two modifications were made to adapt the habituation procedure to the more limited attention spans and exploratory capabilities of infants under the age of 3 months. Since younger infants tend to fuss more easily, show more captive attention, and find it more difficult to look away from a prominent display, the length of "look away" time required to terminate a trial was lowered from 1.5 s used in prior studies with infants of 3 months and older to 1 s for the infants of this study. Further, the maximum number of seconds looking during any particular trial was lowered from 60 to 45 for the infants of this study. This enables the younger infants to disengage more easily and reorient again to initiate the next trial.

A secondary observer monitored visual fixations for 29% of the subjects, 11 of the 20 4-week-olds, 12 of the 28 7-week-olds, and 5 of the 48 11-week-olds. Interobserver reliability was calculated separately for each age group by correlating total fixation time on each trial for observations of the primary and the secondary observer for each infant and then averaging the correlations across infants within an age group. Results indicated Pearson product–moment correlations of .98, ($SD = .02$) for the 4-week-olds, .99 ($SD = .005$) for the 7-week-olds, and .98 ($SD = .03$) for the 11-week-olds.

Results and Discussion

Visual recovery to the two test displays served as the main dependent variable. It was calculated by subtracting the average fixation time across the two (no-change) posthabituation trials from that of the two test trials for each infant. Because only the *relation* between the films and soundtracks was altered for each type of test, and each of the films and soundtracks was familiar to the infant, visual recovery reflects detection of the change in audio–visual relations. The mean visual recovery for infants at each age in the synchrony versus the composition change conditions is depicted in Fig. 1.

Primary analyses. An overall, three-way analysis of variance was first performed on visual recovery with age, condition (synchrony and composition), and stimulus type (orange nuts and yellow washers) as between-subjects factors. Results revealed no main effects of age, condition, or stimulus type (all ps $> .1$), but there was a significant interaction of age and condition [$F(2, 84) = 3.7$, $p = .029$]. To interpret the nature of the interaction effect and to address the primary research question (Did infants detect changes in synchrony and object composition information and at what age?) single sample t tests were performed on the visual recovery scores against the chance value of 0 recovery. Results indi-

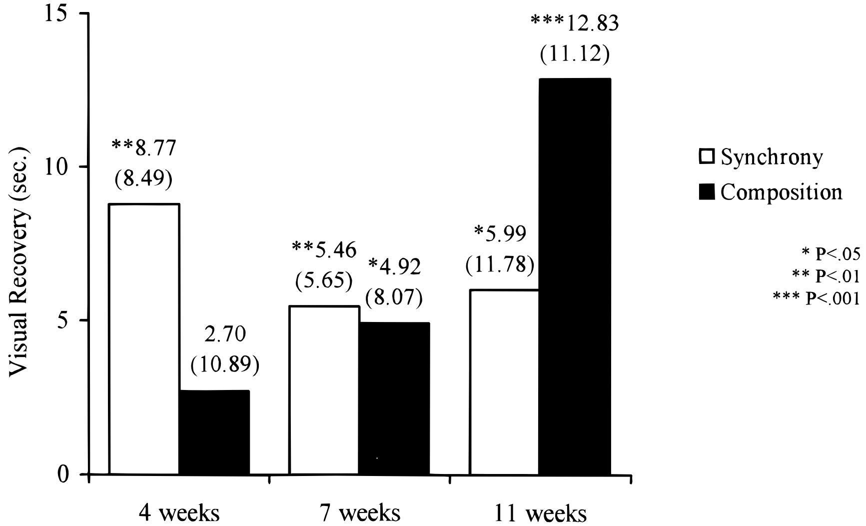

FIG. 1. Mean visual recovery, standard deviations (in parentheses) and significance of visual recovery, according to single sample *t* tests, to the change in temporal synchrony and temporal microstructure specifying object composition at 4, 7, and 11 weeks of age.

cated significant visual recovery to the change in synchrony at all ages [$t(9) = 3.27$, $p = .01$; $t(11) = 3.35$, $p = .007$; $t(23) = 2.49$, $p = .02$, for the 4-, 7-, and 11-week-olds, respectively]. Thus, infants were able to detect the change in audio–visual synchrony relations from habituation to test. However, recovery to the change in information specifying object composition was significant at only the two older ages, but not at 4 weeks [$t(9) = .78$, $p > .1$; $t(15) = 2.44$, $p = .028$; $t(23) = 5.65$, $p = .0001$, for the 4-, 7-, and 11-week-olds, respectively]. In fact, by 11 weeks, visual recovery to the change in composition information was significantly greater than visual recovery to the change in synchrony information [$t(46) = 2.07$, $p = .04$], according to a two-sample *t* test.

To address the second research question (Did infants show a pattern across age and condition consistent with increasing specificity in sensitivity to amodal relations?) analyses of variance and trend analyses were conducted on the visual recovery scores. A two-way analysis of variance with age and condition as between-subjects factors revealed no significant main effects of age [$F(2, 90) = 2.02$, $p > .1$] or condition [$F(1, 90) = .96$, $p > .1$], but a significant interaction between age and condition [$F(2, 90) = 3.23$, $p = .044$] was found, as before. Post hoc analyses were conducted to further evaluate the nature of the interaction effect. They revealed that for the synchrony change, infants showed no significant differences across age. However, for the composition change, visual recovery of the 11-week-olds was greatest and differed significantly from that of both the 7- and 4-week-olds (all p's $< .05$, Tukey HSD). The interaction reflects the finding that mean visual recovery increased across age for the change in composition information but remained relatively constant across age for synchrony information. One-way analyses of variance and trend analysis were also conducted on visual recovery scores for each condition separately to further evaluate the significance of these pattens. They indicated a significant main effect of age for visual recovery to the composition change [$F(2, 47) = 4.72$, $p = .014$] and a significant increasing linear trend across age for sensitivity to composition information[$F(1, 47) = 8.71$, $p = .005$]. However, there was no evidence of a main effect of age for visual recovery to the synchrony change [$F(2, 43) = .364$, $p > .1$] and the linear trend was not significant [$F(1, 43) = .41$, $p > .1$]. Thus, infants showed a significant increase across age in sensitivity to composition information but sensitivity to synchrony information remained stable across age.

Overall, infants demonstrated a pattern of results consistent with an increasing specificity view of perceptual development. Global amodal synchrony relations were detected earliest, by 4 weeks, and remained stable across age, whereas sensitivity to nested temporal relations specifying object composition emerged by 7 weeks and increased with age. In fact, detection of composition information increased linearly with age such that sensitivity to composition was significantly greater than that to synchrony by 11 weeks.

Secondary analyses. Analyses were also conducted to compare the performance of infants across age and condition on the habituation phase. Habituation was examined according to five measures: (a) baseline, defined as the mean length of fixation across the first two habituation trials; (b) mean number of trials needed to reach the habituation criterion; (c) mean number of seconds to habituation summed across trials; (d) mean length of fixation across the two criteria trials; and (e) mean length of fixation across the two no-change, posthabituation trials. Table 1 presents the means and standard deviations for each of these variables at each age.

Analyses of variance were performed separately on each of these variables across age and condition to assess any differences in patterns of habituation. Results indicated no significant effects of condition for any of the variables (all ps $> .05$), indicating that infants who participated in the synchrony versus the composition tests did not differ a priori in their initial or final interest level or patterns of habituation to the displays. Further, as expected, there were significant age differences in patterns of habituation for a number of the measures including baseline [$F(2, 90) = 3.54$, $p = .03$], trials to habituation [$F(2, 90) = 3.68$, $p = .03$], criterion fixation [$F(2, 90) = 4.09$, $p = .02$], and seconds to habituation [$F(2, 90) = 4.87$, $p = .01$]. The main effect of age, however, was not significant for the posthabituation fixation measure [$F(2, 90) = .89$, $p > .1$], indicating that infants did not differ across age in their final level of interest in the displays. Post hoc analyses (Scheffe multiple comparisons, ps $< .05$) were conducted on all significant age effects to specify the nature of the age differences. Analyses for the baseline and number of trials measures indicated that the 11-week-olds showed a greater initial looking time and more trials to habituation than the 4-week-olds

TABLE 1
Experiment 1: Means and Standard Deviations for Five Measures of Habituation for Infants of 4, 7, and 11 Weeks of Age

	Age		
	4 weeks	7 weeks	10 weeks
Baseline[a]			
M	19.56	25.00	28.69
SD	13.81	12.94	12.45
Number of trials			
M	7.90	9.39	7.88
SD	2.38	2.69	2.29
Number of seconds			
M	96.53	153.53	141.66
SD	39.05	67.06	70.89
Criterion looking[b]			
M	3.40	5.62	6.28
SD	2.80	4.11	3.91
Posthabituation looking[c]			
M	8.95	7.39	7.08
SD	5.64	5.59	5.21

[a] Baseline is the mean length of fixation across the first two habituation trials.
[b] Criterion looking is the mean length of fixation across the two habituation criterion trials.
[c] Posthabituation looking is the mean length of fixation across the two no-change posthabitation trials.

($ps < .05$). The analysis on the number of seconds to habituation indicated that the 4-week-olds spent the least time overall looking at the displays; less than either the 7-week-olds or the 11-week-olds. Further, the 4-week-olds spent less time during the criterion looking trials than did the 11-week-olds. Thus, in general, 4-week-olds looked least initially and on the criterion trials and took the least amount of time to habituate. The lower level of overall looking is noteworthy given that younger infants are typically less efficient processors than older infants or children (Gibson, 1969). Given that infants of all ages viewed the stimuli in an infant-controlled procedure and attended to the displays until they reached habituation, it is reasonable to assume that the attentional differences were infant-controlled and a result of differential interest and attentional patterns across age. The lower level of looking by 4-week-olds may reflect their more limited attention spans and result from a failure to notice or attend to embedded amodal relations. This fits well with the finding that infants of this age showed no evidence of detecting audio–visual relations based on embedded amodal temporal structure. In contrast, by 7 weeks, when detection of nested amodal relations was evident, infants showed a clear increase in attention to the displays. Overall amount of looking time (seconds to habituation) was significantly greater by 7 weeks and equaled that of the 11-week-olds. Thus, it may be that the increase in attention to the displays between 4 and 7 weeks is a result of the 7-week-olds discovering a new level of audio–visual stimulation to explore. That is, it may reflect the shift from processing of global to nested amodal relations.

Given that infants differed across age in their overall processing time of the audio–visual events, an analysis of covariance was performed to determine to what extent this variable impacted the observed differences in visual recovery times. That is, after the visual recovery scores were adjusted for differences in processing time, would the interaction between age and condition still reach significance? Results of the analysis indicated that the seconds to habituation measure was not a significant predictor of visual recovery [$F(1, 89) = .52$, $p > .1$] and that the interaction between age and condition was still significant when seconds to habituation was used as a covariate [$F(2, 89) = 3.134$, $p = .048$]. Further, post hoc tests were conducted on the adjusted means and results revealed the same pattern as with the unadjusted means. That is, there were no significant differences across age in visual recovery to the synchrony change; however, for the composition change, visual recovery of the 11-week-olds differed significantly from that of the 7- and 4-week-olds (p's $< .01$, Tukey HSD). Thus, the differences in processing time do not significantly impact the differences in visual recovery.

Secondary analyses were also performed to assess any effects of stimulus pair on visual recovery scores. An analysis of variance with age and stimulus pair (orange nuts vs yellow washers) as main factors was conducted and revealed no significant effect of stimulus pair [$F(1, 90) = 1.28$, $p > .1$] or interaction of stimulus pair and age [$F(2, 90) = .24$, $p > .1$].

GENERAL DISCUSSION

The world of multimodal events makes global and nested audio–visual relations available. When infants encounter natural, dynamic events, they appear to actively abstract some properties and ignore others. The present research suggests that this pattern of selectivity may be lawfully organized and developmentally regulated in the first months of life. Results of the present research demonstrate a pattern of perceptual differentiation consistent with an increasing specificity view of development. Infants were able to detect global temporal synchrony relations by 4 weeks of age, whereas detection of audio–visual relations specifying object composition was evident only by 7 weeks. These findings suggest that in the domain of intersensory perception, infants appear to abstract global information developmentally prior to nested relations. Further, the present findings reveal different developmental trajectories for sensitivity to synchrony versus composition information. That is, detection of audio–visual synchrony was evident by 4 weeks and remained constant across age. In contrast, detection of audio–visual information specifying object composition was not apparent until 7 weeks and showed a significant linear increase across age. By 11 weeks sensitivity to composition relations was significantly greater than sensitivity to synchrony relations. Thus, infants showed an increasing sensitivity to audio–visual composition information across age, whereas sensitivity to synchrony was already present by 4 weeks and remained stable across age. This pattern illustrates increasing specificity across development in processing of global to nested relations and in perceptual learning of nested relations. In this manner infant perception can become increasingly aligned with the nature of stimulation in the environment across development.

The present study also adds to the evidence accruing that suggests that sensitivity to temporal synchrony in naturalistic audio–visual events emerges very early in infancy and likely serves as one of the first and most important bases for organizing perceptual development (see also Bahrick & Pickens, 1994; Lewkowicz, 1989, 2000). By detecting synchrony, infants are able to determine which sights and sounds belong together and constitute unitary events and which are unrelated. The present findings revealed that synchrony detection was present at 4 weeks, the earliest age tested, and the possibility thus remains that sensitivity to synchrony may even guide intersensory perception at birth. Recent findings (Slater, Quinn, Brown, & Hayes, 1999) are consistent with this view and show sensitivity to synchrony between the infant's own actions and environmental consequences within a few days of birth.

Further, the present study also reveals detection of amodal information specifying object composition at a younger age than previously noted. Prior research (Bahrick, 1992) had demonstrated infant sensitivity to this information at 15–16 weeks. The present study reveals the emergence of this ability at 7 weeks and a dramatic increase in sensitivity to these relations by 11 weeks, when this sensitivity guides visual exploration and overshadows interest in synchrony relations. It may be that once infants perceive which sights and sounds belong together by detecting synchrony, they show greater interest in more specific, nested aspects of the events such as their composition or substance.

The present empirical findings are consistent with proposals from a variety of investigators derived from reviews of the diverse literature in the area of intersensory development. Walker-Andrews (1997) concluded from a comprehensive review of the literature on the development of face and voice perception that recognition of emotional expressions develops first in the context of multimodal information with the dynamic face and voice of a speaking person. Then, the emotional properties of the face and voice become progressively differentiated and discriminated separate from the whole. This demonstrates increasing specificity in the perception of affective information. Similarly, Lewkowicz (1989, 2000) proposed a developmental model for the detection of intersensory temporal relations consistent with the principle of increasing specificity. He argued that responsiveness to four basic properties of audible and visible events, synchrony, duration, rate, and rhythm, emerge in a sequential, hierarchical fashion. Infants initially detect synchrony, the most basic temporal relation, and this enables them to detect the others in order of increasing complexity. Other investigators have also argued that perception is at first global and only later becomes more specific and differentiated. For example, according to the "intensity matching" view

(Lewkowicz & Turkewitz, 1980; Schnierla, 1957; Turkewitz, Gardner, & Lewkowicz, 1984) infants initially do not differentiate information from the different senses. Rather, they respond to the overall intensity of stimuli summed across the modalities. Similarly, Bower (1974) and Maurer (1993) also argued that infants begin life with undifferentiated senses and through development, the senses and the information conveyed through them become gradually differentiated. These varied proposals, consistent with the increasing specificity view, are based on reviews of the evidence from a diversity of approaches with methods, stimuli, and tasks that vary widely.

The present study, taken together with previous empirical work on this topic (Bahrick 1992, 1994; Gogate & Bahrick, 1998; Hernandez-Reif & Bahrick, 2001), provides converging evidence for a pattern of increasing specificity across development. Infants apparently detect global, amodal synchrony relations earliest, then nested amodal relations, and finally arbitrary, modality-specific, audio–visual relations within a domain. There appears to be a developmental lag between detection of global amodal, nested amodal, and arbitrary audio–visual relations. Findings of the present study revealed that infants detected the global, amodal relations by 4 weeks and the nested, amodal relations at 7 weeks. However, it was not until 7 months of age that infants showed evidence of detecting the arbitrary audio–visual relations between the pitch of the impact sound and the color/shape of these same objects (Bahrick, 1994). This developmental lag is adaptive because it can foster learning about consistencies and regularities across the senses that are context-free prior to learning about audio–visual relations that are idiosyncratic and vary from one context or event to another.

Further, research has found that detection of global audio–visual relations can guide and constrain learning about nested relations. For example, infants detect the arbitrary relation between a speech sound and an object only when the object and sounds are presented in synchrony and not when they are out of synchrony or the object is still (Gogate, 2001; Gogate & Bahrick, 1998). The synchrony highlights the relatedness between the sound and the object and promotes further processing of the nature of this relation. Hernandez-Reif and Bahrick (2001) have shown that 6-month-old infants detect the arbitrary relation between the tactually perceived shape of an object and a specific color/pattern only when amodal information for object shape unites their visual and tactile exploration. Further, Bahrick (1988) demonstrated that infants only learned to relate a film of a brightly colored, moving object and a distinctive soundtrack if the object motions and sounds shared both a temporal synchrony relation and a common temporal microstructure specifying object composition. They were unable to learn if either synchrony or composition information was incongruent. Recently Bahrick and Lickliter (2000) demonstrated that 5-month-olds could detect amodal rhythm only when it was presented bimodally rather than unimodally, and only when the bimodal presentation was temporally synchronous. Thus, detection of global amodal relations appears to guide and constrain detection of nested amodal and arbitrary audio–visual relations. These findings suggest that the pattern of increasing specificity not only characterizes a developmental sequence, but as a result, also characterizes a sequence of processing within an episode of exploration. That is, exploration of an object or event may first progress through a phase in which global properties are processed. Then, with additional exploration time, nested relations may be differentiated, and finally, if additional time is available, arbitrary relations may be abstracted. As infants become more skilled at abstracting global, amodal relations through development, it seems likely they would do so with increasingly less processing time, allowing more processing time for nested relations. Direct empirical testing of this notion must be undertaken before it can be substantiated.

In sum, the present study, together with those of Bahrick (1992, 1994) provided direct evidence for the principle of increasing specificity in the development of audio–visual event perception. These related studies demonstrate developmental trends in perception using methods, tasks, and stimuli that are consistent across studies and age. Further, they tested a priori hypotheses that were generated from the principle of increasing specificity and were based on a priori definitions of global and embedded audio–visual relations. As a whole, they provide a systematic body of evidence for a general developmental principle of increasing specificity. Further evidence from other areas of development will be important in evaluating the generalizability of this principle across domains.

Detection of object and event properties in order of increasing specificity is adaptive. Natural events in the infant's environment typically make both global and specific information available. Early sensitivity to global information at the expense of specific information can serve to focus infant attention on meaningful, unitary audio–visual relations, appropriately guide and constrain subsequent exploration, and buffer infants from learning inappropriate, incongruent relations across the senses. Perceptual development in order of increasing specificity can thus provide a fundamental, organizational framework for the development of veridical perception.

REFERENCES

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development*, **6,** 429–451.

Bahrick, L. E. (1987). Infants' intermodal perception of two levels os temporal structure in natural events. *Infant Behavior and Development*, **10,** 387–416.

Bahrick, 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,** 197–207.

Bahrick, L. E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio–visual relations. *Journal of Experimental Child Psychology*, **53,** 180–199.

Bahrick, L. E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology*, **6,** 111–123.

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. (1994). Amodal relations: The basis for intermodal perception and learning. In D. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 205–233). Hillsdale, NJ: Erlbaum.

Barrera, M. E., & Maurer D. (1981). Recognition of mother's photographed face by the 3-month old infant. *Child Development*, **52,** 714–716.

Bertenthal, B. I., Haith, M. M., & Campos, J. J. (1983). The partial-lag design: A method for controlling spontaneous regression in the infant-control habituation paradigm. *Infant Behavior and Development*, **6,** 331–338.

Bornstein, M. H. (1978). Chromatic vision in infancy. In W. H. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 12, pp. 117–182). New York: Academic Press.

Bower, T. G. R. (1974). *Development in infancy*. San Francisco: Freeman.

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. (1984). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of the Acoustical Society of America*, **98,** 1372–1379.

Clarkson, M. G., Martin, R. L., & Miciek, S. G. (1996). Infants' perception of pitch: Number of harmonics. *Infant Behavior and Development*, **19,** 191–197.

Dannemiller, J. L. (1989). A test of color constancy in 9- and 20-week-old human infants following simulated illuminant changes. *Developmental Psychology*, **25,** 171–184.

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

DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: Preference, discrimination, and reinforcing value. *Developmental Psychobiology*, **17,** 481–41.

Fagan, J. F. (1972). Infants' recognition memory for faces. *Journal of Experimental Child Psychology*, **14**, 453–476.

Fagan, J. F. (1976). Infant's recognition of invariant features of faces. *Child Development*, **47,** 627–638.

Fantz, R. L., Fagan, J. F., & Miranda, S. B. (1975). Early visual selectivity as a function of pattern variables, previous exposure, age from birth and conception, and expected cognitive deficit. In L. P. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition: Basic visual proceses* (vol.1, pp. 249–345). New York: Academic Press.

Fenwick, K. D., & Morrongiello, B. A. (1998). Spatial co-location and infants' learning of auditory-visual associations. *Infant Behavior and Development*, **21,** 745–760.

Fernandez, M. & Bahrick, L. E. (1994). Infants' sensitivity to arbitrary object–odor pairings. *Infant Behavior and Development*, **17,** 471–474.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton–Century–Crofts.

Gogate, L. J. (2001). *Intersensory redundancy facilitates learning of arbitrary minimal pair-object relations by 8- but not 7-month-old infants: Evidence for a dynamic system*. Manuscript submitted for publication.

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,** 1–17.

Hernandez-Reif, M., & Bahrick, L. E. (2001). The development of visual-tactual perception of objects: Amodal relations provide the basis for learning arbitrary relations. *Infancy*, **2,** 51–72.

Lewkowicz, D. J. (1989). The role of temporal factors in infant behavior and development. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 9–62). Amsterdam, The Netherlands: North-Holland.

Lewkowicz, D. J. (1992). Infants' response to temporally based intersensory equivalence: The effect of synchronous sounds on visual preferences for moving stimuli. *Infant Behavior and Development*, **15,** 297–324.

Lewkowicz, D. J. (1996). Perception of auditory–visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance*, **22,** 1094–1106.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, **126,** 281–308.

Lewkowicz, D. J., & Lickliter, R. (1994). *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum.

Lewkowicz, D. J., & Turkewicz, G. (1980). Cross-modal equivalence in early infancy: Auditory–visual intensity matching. *Developmental Psychology,* **16,** 597–606.

Maurer, D. (1993). Neonatal synesthesia: Implications for the processing of speech and faces. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 109–124). Boston: Kluwer.

Mendelson, M. J., & Ferland, M. B. (1982). Auditory–visual transfer in four-month-old infants. *Child Development,* **53,** 1022–1027.

Morrongiello, B. A., Fenwick, K., & Nutley, T. (1998). Developmental changes in associations between auditory–visual events. *Infant Behavior and Development*, **21,** 613–626.

Nelson, C. A., & Horowitz, F. D. (1983). The perception of facial expressions and stimulus motion by 2- and 5-month-old infants using holograhic stimuli. *Child Development*, **54,** 868–877.

Peeples, D. R., & Teller, D. Y. (1975). Color vision and brightness discrimination in 2-month-old human infants. *Science*, **189**(4208), 1102–1103.

Pickens, J. (1994). Perception of bimodal distance relations by 5-month-old infants. *Developmental Psychology*, **30,** 537–544.

Reardon, P., & Bushnell, E., W. (1988). Infants' sensitivity to arbitrary pairings of color and taste. *Infant Behavior and Development,* **11,** 245–250.

Schnierla, T. C. (1957). The concept of development in comparative psychology. In D. B. Harris (Ed.), *The concept of development*. Minneapolis: Univ. of Minnesota Press.

Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory–visual pairings. *Developmental Science,* **2,** 333–338.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, **15,** 626–636.

Teller, D. Y., & Bornstein, M. H. (1987). Infant color vision and color perception. In P. Salapatek & L. Cohen (Eds.), *Handbook of infant perception: From sensation to perception* (Vol. 1, pp. 185–236). Orlando, FL: Academic Press.

Turkewicz, G., Gardner, J. M., & Lewkowicz, D. J. (1984). Sensory;shperceptual functioning during early infancy: The implication for a quantitative basis of responding. In G. Greenberg & E. Tobach (Eds.), *Behavioral evolution and integration levels* (pp. 167–195). Hillsdale, NJ: Erlbaum.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, **33,** 514–535.

Walker-Andrews, A. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, **121**(3), 437–456.

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infant's bimodal perception of gender. *Ecological Psychology,* **3,** 55–75.

Walker-Andrews, A. S., & Lennon, E. M. (1985). Auditory–visual perception of changing distance by human infants. *Child Development*, **56,** 544–548.

Received August 13, 1999; revised June 19, 2000; published online May 8, 2001