INFANT BEHAVIOR AND DEVELOPMENT 6, 429–451 (1983)

Infants' Perception of Substance and Temporal Synchrony in Multimodal Events*

LORRAINE E. BAHRICK
Florida International University, Miami

Four-month-old infants can detect a relationship between the soundtracks and films of rigid and elastic objects in motion. When movies of these two kinds of objects were shown side by side along with the soundtrack to one of them, infants predominately watched the sound films. Three subsequent studies explored the basis of this intermodal perception. It was found that infants detected bimodal temporal structure specifying object rigidity and elasticity, and temporal synchrony between the sights and sounds of object impact. These results support the view that the development of intermodal perception is based on the detection of invariant relations in visual and acoustic stimulation.

intermodal perception intersensory perception amodal perception substance perception synchrony perception invariant detection

Most events are multimodal. That is, they make information available to several sensory systems simultaneously. The adult perceives a stable world of unitary objects and events through a continuously changing flux of multimodal stimulation. The sound of a friend's voice along with the sight of his/her moving face is perceived as a single, unitary event; that of the friend speaking. It is automatically differentiated from the sight of another passing person or the sound of a nearby telephone ringing. The adult perceiver has developed intermodal knowledge and expectations about events such that information picked up through one modality elicits expectations about what is to be perceived through another modality.

How does the infant develop the capabilities of the adult perceiver? More specifically, how does he determine which patterns of auditory and visual stimulation constitute single, unitary events, and which are unrelated to one another? Recent research has suggested that the neonate explores his environment through spatially coordinated systems of audition and vision. Wertheimer (1961) and, more recently, Muir and Field (1979) report that the neonate has a tendency to direct his eyes toward the locus of a sound. This coordination provides an ideal opportunity for learning about multimodal events. It maximizes the likelihood of picking up

* This paper is based on a doctoral dissertation submitted by the author to Cornell University. The research was supported in part by a National Institute of Mental Health Training Grant, 5 T01 MH08520-15, to the Department of Psychology, Cornell University, a National Science Foundation grant, BNS76-14942, awarded to Eleanor J. Gibson, a grant to Eleanor J. Gibson from the Spencer Foundation, and a grant from the National Institute of Child Health and Human Development, #1 R01 HD15275-01. Correspondence and requests for reprints should be addressed to: Lorraine E. Bahrick, Department of Psychology, Tamiami Campus, Florida International University, Miami, FL 33199.

visual information about acoustically detected events. However, there are always many simultaneously visible objects within one direction of gaze. How can the infant determine which of the many visible objects, if any, is the source of the sound he is hearing?

According to Gibson (1969), invariant information unites multimodal stimulation from a single event. Some properties of objects are specified to several perceptual systems. When invariants are detected that specify the same properties in different modes, a unitary object is perceived. For example, since events occur over time, they may be characterized by a temporal structure; temporal synchrony, tempo of action, or temporal pattern can be invariant over the visual and acoustic portions of an event. The same temporal structure may be picked up through both modalities. The temporal structure is termed *amodal* since it is neither specific to vision nor audition, but common to both modes. For example, when two hands clap out a rhythm, the same rhythm can be both seen and heard. The detection of this invariant structure would seem an ideal way for the infant to develop intermodal knowledge. Because the same structure may be picked up through two modalities simultaneously, it can (a) specify the unitary nature of the audiovisual event, and (b) separate it from other co-occurring events that do not share its structure. Sensitivity to this kind of invariant should enable the infant to readily determine which visible object is part of the event he is hearing.

Recent research has just begun to delineate some of the audiovisual invariants young infants are able to detect. Spelke (1979) found that 4-month-olds were sensitive to temporal synchrony and tempo of action common to filmed events and their soundtracks. Infants can also detect simple rhythmic patterns across visual and acoustic presentations even when the patterns are presented successively (Allen, Walker, Symonds, & Marcell, 1977).

The research reported here investigated infants' sensitivity to two kinds of temporal invariants in naturalistic, audiovisual events: (a) substance information, in this case temporal information specifying the rigidity or elasticity of an object in motion, and (b) temporal synchrony information, here the coincidence between the sights and sounds of impact of an object striking a surface.

Four studies were conducted to explore these issues. Stimulus films were designed to portray two types of events; hard objects undergoing rigid perspective transformations (blocks banging) and soft objects undergoing deformations (sponges squashing). In Experiment 1, subjects were shown the two films side by side along with the appropriate and synchronized soundtrack to one of them. Infants showed a looking preference for the sound films. Experiments 2, 3, and 4 investigated the basis of this preference. Experiment 2 placed synchrony and substance information in conflict with one another. Infants viewed the two films side by side along with a soundtrack that was synchronized with one film but appropriate to the substance of the other. Experiment 3 examined infants' detection of information specifying object substance in the absence of synchrony information. Experiment 4 investigated infants' detection of temporal synchrony information in the absence of substance information. Infants were found to be sensitive to both synchrony and substance information.

EXPERIMENT 1

This study examined whether infants could detect a relationship between the auditory and visual components of complex, naturalistic events. It tested the appropriateness of the present stimulus materials and provided a point of departure for the three subsequent studies that investigated the basis of intermodal perception. Previous research had already demonstrated intermodal perception of naturalistic events by infants (Bahrick & Walker, 1978; Spelke, 1976). When infants viewed two filmed episodes side by side along with the soundtrack to one of them, quite robust-looking preference for the sound films were obtained. However, the basis for these effects was not systematically investigated with these materials. Spelke (1979) later examined infants' detection of temporal invariants using simpler, repetitive events. She found infants' preferences to be attenuated with these materials, but was able to demonstrate their sensitivity to temporal invariants using a different "search" procedure. It seemed possible that the richness and complexity of the natural events was responsible for the sizable effects in the earlier studies. The present experiment thus used complex, unpredictable events to replicate Spelke (1979) and Bahrick and Walker (1978).

The study incorporated both aspects of Spelke's (1979) procedure; the "preference" and "search" phases. During the preference phase, subjects viewed the two films side by side, with the soundtrack to one of them coming from a centrally placed speaker. They then viewed the films for an additional period along with the other soundtrack. It was expected that subjects would preferentially fixate the sound specified films, demonstrating a sensitivity to audiovisual correspondence. A search phase always followed the preference phase and was identical for all studies. It was included as a means of determining what the subject had learned about the events during the preference phase and served as a basis of comparison among the three studies. In the search phase, subjects again viewed the two films side by side while intermittent 5-sec presentations of the soundtracks from both films were interspersed. It was expected that subjects would "search" first in the direction of the sound specified films if they had acquired intermodal knowledge about the events during the preference phase.

Method

Subjects. Twenty 4½-month-old infants, 10 males and 10 females, participated. Their mean age was 143.8 days with a range of 129 to 157 days. Subjects were located through birth announcements in local newspapers. Three additional subjects were tested and excluded from the experiment: one, because of experimental error, and two, on the basis of age. It was found that younger infants often demonstrated strong side biases as their heads fell to one side, constraining visual exploration. A lower age limit of 127 days was thus set for subjects in all studies.

Stimulus materials. Three-minute, 15-sec, Super-8, sound, color films were made of two kinds of episodes depicted in Fig. 1. The film of the rigid object showed two yellow wooden blocks striking each other or a wooden box in an erratic

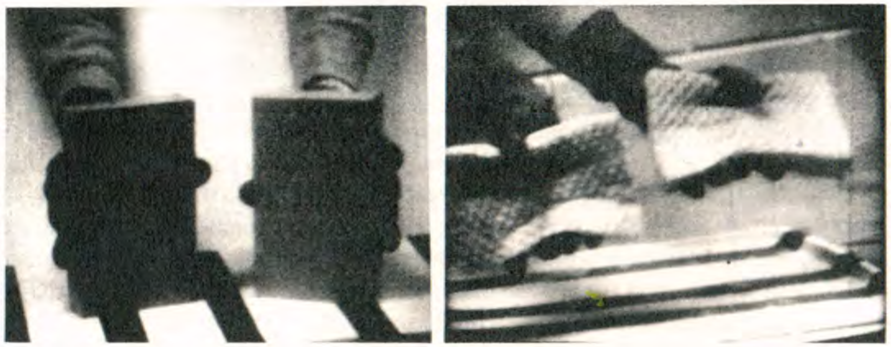

Figure 1. Photograph of the two films used as stimulus materials in Experiments 1–4.

pattern. The motions were primarily rigid perspective transformations, and the sounds produced by their impact were abrupt, sharp banging sounds. The film of the elastic object depicted two yellow water-soaked sponges being abruptly squashed against each other or singly in an erratic pattern. Their motions were primarily deformations, and they produced squashing sounds that were softer and somewhat more prolonged than the sounds of the blocks. The soundtracks to both episodes were recorded live. Both events were characterized by erratic and unpredictable rhythmic and spatial patterns, yet displayed similar rates of motion when averaged over the entire length of the films. On the average impact sounds occurred approximately 4 times every 5 sec, but from one moment to the next the pattern was variable and unpredictable. These shared properties ensured that the infants would not be able to discriminate between the films on the basis of tempo of action or rhythmic pattern. Precautions were also taken to equate color and amount of motion. The sponges and blocks were both bright yellow and of approximately the same size (1 × 3 × 6 in). Two hands wearing blue rubber gloves manipulated the objects in both films, and the surfaces beneath them were white with bright blue stripes. The movements of the objects in both events contained vertical, horizontal, and diagonal components.

Apparatus. Infants were seated in an infant seat in a booth enclosed by curtains on two sides. A rear projection screen of translucent glass (50 × 80 cm) faced the infant about 60 cm from the subject's eyes. The two films were rear projected side by side using two Bolex (SP 80) sound movie projectors. The two images were approximately 30 × 33 cm each and were separated by a 10-cm wooden strip which held a flashlight. The soundtrack to one of the films was always played through a speaker centered midway between the two images and about 1.1 meters behind the screen.

An observer, blind to the lateral positions of the films, monitored the subject's fixations from a peep hole just below the projection screen. He depressed one button when the subject fixated the right-hand image and another when the subject fixated

the left-hand image. Neither button was depressed when the subject fixated neither film. The fixations were recorded with a Harvard Apparatus event recorder along with the soundtracks to each event.

Procedure (preference phase). Each subject viewed two films, one of the blocks and the other of the sponges, side by side for 90 sec, along with the soundtrack to one of them coming from the centrally placed speaker (viewing 1). The two films were then projected for an additional 90 sec along with the other soundtrack (viewing 2). The order of soundtrack presentation and lateral positions of the films were counterbalanced across subjects. Ten subjects received the soundtrack to the blocks film first and the sponges second, while the other 10 heard the soundtracks in the reverse order. Half the subjects in each group saw the film of the blocks on the right and the sponges on the left, while the other half received the opposite arrangement.

Procedure (search phase). The search phase always followed the preference phase after a 2–3 minute break and was identical across the 4 experiments. The two films were continuously projected side by side for an additional 3 minutes. They were accompanied by intermittent 5-sec presentations of the natural, synchronized soundtracks from the two films in a randomized pattern. The measure of interest was the direction of the subject's first look at the onset of each soundtrack presentation. Because of this, it was necessary to ensure that the subject was not already fixating one of the films when the soundtracks were turned on. Consequently, the infant's visual behavior determined the onset of each trial. A flashlight positioned between the two films was blinked off and on to attract the infant's attention to center screen. A soundtrack was turned on only when the infant was fixating neither film.[1] Subjects thus received different numbers of trials depending upon their visual behavior. Some subjects received as few a 6 trials while others received up to 15 usable trials. A criterion of at least 5 usable trials was preset for subjects to be included in this part of the experiment. Only one subject was excluded for this reason.

The observer monitored the subject's fixations during this phase of the study using two sets of buttons; right/left looking was recorded with one set of buttons when the films were silent and with the other when the soundtracks were audible. The observer was again blind to the lateral positions of the films.

The order of soundtrack presentation was semirandom with the restriction that the same soundtrack was played no more than twice in succession. The lateral positions of the two films remained the same for a given infant throughout the preference and search phases of the experiment. This provided the infant with the opportunity for learning about the relative locations of the two films during the preference phase and the possibility of using this information to glance appropriately during the search phase.

[1] Despite procedural precautions, subjects occasionally managed to shift fixation from center screen to one of the films before E could begin a trial by turning on a soundtrack. Such trials were detected from the event recorder output and discarded.

Interobserver reliability was calculated for the primary observer with a second trained observer for both the visual search and preference phases of all four studies. Reliabilities for the preference phase were calculated on the basis of the number of seconds for which both observers agreed on the direction of the subject's fixation, out of the total number of seconds the films were shown. For any 1-sec interval, a disagreement was scored if observers showed an inconsistency of .2 sec or more. Interobserver reliability for the preference phase averaged .93 and was calculated on the basis of 15 subjects out of the total of 80 across the four experiments. Reliabilities for the search phase were derived from the number of agreements between the two observers on the direction of the subject's first look after the onset of each soundtrack, out of the total number of usable trials. Reliability averaged .86. This was based on only 12 of the subjects since three could not complete the search phase because of excessive fussiness.

Thus, this study consisted of two phases, the preference and the search phase. During viewing 1 of the preference phase the two films were projected side by side for 90 sec along with the soundtrack to one of them. During viewing 2 the films were projected along with the soundtrack to the other one. After this preference phase, the search phase followed. The two films were projected for an additional 3 min along with intermittent presentations of the soundtracks from each.

Results and Discussion

Results of the experiment are shown in Table 1. Looking preferences for each infant were derived separately from the two viewing periods. The time spent looking to the sound film was expressed as a proportion of the total time spent looking to both films. T-tests comparing the mean looking preferences for the sound-specified films against a chance value of .5 were performed. (All p-values reported in this study and the subsequent ones are two-tailed.) During the first viewing of the preference phase infants spent a significantly greater proportion of the time, .603, viewing the sound-specified films, $t(19) = 2.56$, $p < .02$.[2] Fifteen of the 20 subjects demonstrated this preference effect, spending a majority of the time fixating the sound-specified films.

During the second viewing of the preference phase, infants demonstrated no significant preference for either film. This is consistent with findings of Spelke (1979) using this procedure and with Watson's (1968) analysis of infants' fixation behavior under similar visual choice conditions. He found that infants demonstrated an acquired position response. If they had been fixating a display on the right side, they began the subsequent trial with a dominant tendency to fixate that side even though stimulus conditions had changed.

Since several infants became restless and fussy toward the end of the session, it was hoped that a shorter preference phase of two 60-sec viewings would be sufficient in the subsequent studies. Results of the preference phase based on the

[2] Reported degrees of freedom for t-tests are adjusted values given by a computer program (Minitab; Ryan, Joiner, & Ryan, 1976) which provides more conservative tests of significance to compensate for heterogeneity of variance.

TABLE 1
Experiment 1: Looking Times, Preferences, and Search Proportions for the Sound-Specified Films

	Viewing 1 (N = 20)		*Viewing 2 (N = 20)*		*Search (N = 17)*[a]
	90 sec	*60 sec*[d]	*90 sec*	*60 sec*	
Mean looking time (sec)					
Sound film	46.5	31.8	29.5	20.0	
Silent film	30.4	20.3	39.8	27.1	
Total	76.9	52.1	69.3	47.1	
Mean preference[b]					
Proportion	.60	.61	.43	.43	.57
SD	.18	.19	.27	.26	.12
t	2.56**[c]	2.69**[c]	1.13	1.29	2.40*[c]

[a] These figures are based on the proportion of first looks to the sound-specified films.
[b] Looking preferences were computed by dividing the number of seconds each subject looked at the sound film by the total time spent looking at both films, and averaging over the 20 values. These average proportions differ from the proportions derived from "Sound film" to "Total" above them because the mean of ratios is not, in general, equal to the ratio of means.
[c] p values are two-tailed; $*p < .05$; $**p < .02$.
[d] This column displays results of analyses performed on only the first 60 seconds of the 90-sec viewing period.

first 60 sec of each 90-sec viewing period were calculated and were essentially identical to those obtained for the full viewing period (see Table 1). In addition, two analyses of variance were conducted to determine whether subjects displayed any trends over time in looking to the sound appropriate films during the first or second viewing of the preference phase. Each 90-sec viewing period was divided into 30-sec segments and preferences were derived separately for each segment. No trends over time within viewing period 1 or 2 were found, $F(2,38) = .868$, $p > .05$, $F(2,38) = .169$, $p > .05$, respectively. Consequently, a preference phase of two 60-sec viewing periods was adopted for use in the later studies.

A two-way analysis of variance indicated no significant side or film preferences during the first viewing of the preference phase, $F(1,16) = .589$, $p > .05$, $F(1,16) = 1.930$, $p > .05$, respectively. The interaction between side and film preference was also insignificant, $F(1,16) = .062$, $p > .05$.

Results of the search phase were expressed as a proportion of the trials on which the infant looked first to the sound-specified films out of the total number of trials received. No trials on which the infant had already been looking to a film when the sound began were included. Seventeen of the 20 subjects participated in the search phase; the remaining 3 subjects were excluded because of excessive fussiness. The proportion of first looks directed toward the sound specified films averaged .572. This was significant when compared against the chance value of .5, $t(16) = 2.397$, $p < .05$.

Subjects demonstrated a significant preference for the sound-specified films during the preference phase and looked first more often to the sound films during the search phase. These results support those of Bahrick and Walker (1978) and Spelke (1976), demonstrating intermodal exploration of naturalistic events by young infants. Infants were apparently able to detect information common to the films and their soundtracks.

What kind of information might the infants have detected? They could not have relied upon sound localization since the soundtracks came from centrally placed speakers. Nor could they have used tempo of action or rhythmic pattern since the films shared a common tempo and had an unpredictable rhythmic pattern. Two kinds of temporal invariants, however, did unite the films with their soundtracks, and could have been detected by infants. First, there was synchrony between the sights and sounds of impact in these events. Every time an object was seen to contact a surface, a sound was heard at exactly that moment. Second, there was temporal information common to the visible motions and sounds of the objects specifying their rigidity or elasticity. The time courses and types of motions and sounds of a soft object hitting a surface may be similar, and they may differ markedly from the motions and sounds of a hard object. Objects made of hard substances move in a rigid way and produce abrupt, sharp sounds upon impact. Those of an elastic substance may move in a deforming manner and produce softer, more prolonged sounds upon impact. The temporal information available both visually and acoustically specifies the object's substance. Experiment 1 does not indicate whether infants picked up this kind of "substance" information, temporal synchrony, or both. Experiments 2, 3, and 4 were conducted to determine this.

EXPERIMENT 2

In the previous study infants looked to the sounding object when the sound was both synchronous with its movements and appropriate to its substance. The present study investigates whether infants will look preferentially to a sound-synchronized object even when the sound is inappropriate to its substance. Synchrony and substance information were placed in conflict with one another, so that the soundtrack was in synchrony with the objects in one film and appropriate to the substance of those in the other film. Thus, this study was identical to Experiment 1, except that the soundtracks were synchronized with the wrong type of object. The sounds of the sponges squashing were dubbed onto the film of the blocks, while the sounds of the blocks banging were dubbed onto the film of the sponges. With this design one could assess whether infants perceived an intermodal relationship between the objects and sounds on the basis of synchrony, substance information, or both. It was expected that infants would detect a relationship between a film and a soundtrack, only if the sounds were synchronized and appropriate to the substance of the objects, as in Experiment 1. In this study, looking to the sound-synchronized film should be attenuated due to the inappropriate nature of the sounds if infants are indeed sensitive to bimodal information specifying object substance.

There are some findings which lend tentative support to this hypothesis. In a series of intermodal preference studies conducted by Spelke (1978), Moog synthesizer sounds were synchronized with the motions of toy animals bouncing up and down. Infants showed little or no preference for the sound-synchronized events. Although there was a tendency for infants to develop a preference for the synchronized film during the latter half of the viewing period, these preferences were much weaker than those obtained in later studies when the synthesized sounds were replaced by natural impact sounds. The natural impact sounds were more appropriate to the impacting motions of the toy animals than the synthesized sounds. These findings suggest that infants may be sensitive to the relationship between the type of sound and the type of visual impact of an object. Experiment 2 tests this possibility more directly. In sum, it consisted of two parts, a visual preference phase and a search phase. The preference phase was again comprised of two viewing periods and the soundtracks were synchronous but inappropriate to the substance of the objects. The search phase was identical to that of Experiment 1, and tested what the infant had learned during the preference phase.

Method

Twenty subjects, 9 males and 11 females, participated in the experiment. Their mean age was 137.3 days, with a range of 128 to 154 days. One additional subject was run and excluded from the study due to excessive crying. Since this study was designed to be directly compared with Experiment 1, the two groups of subjects were run concurrently.

The stimulus films were the same as those used in Experiment 1. However, new soundtracks were dubbed onto the films so that the sounds of the sponges coincided with the visual impacts of the blocks, and the sounds of the blocks coincided with the visual deformations of the sponges. This was accomplished by recording the new inappropriate sounds directly onto the magnetic sound strips of the films through the movie projector. While viewing the projected film, the experimenter recreated the new sounds in synchrony with each visible impact of the projected object, and recorded them onto the film. Any errors in synchrony were corrected by simply rerecording. The final product considted of two films with dubbed soundtracks in very close synchrony with the visual impacts of the objects.

The apparatus and procedures were the same as those used in the first study. Subjects viewed two films, the blocks and the sponges, side by side for 60 sec, along with an inappropriate soundtrack synchronized with one of them. During the second 60-sec viewing, the other film was synchronized with its inappropriate soundtrack. Thus, during each viewing the sounds were synchronized with the motions of the objects in one film and appropriate to the substance of the objects in the other film. The lateral positions of the films and order of soundtrack presentation were counterbalanced across infants as in Experiment 1.

The search phase followed the preference phase as before. The search phase was identical to that of Experiment 1. That is, the original nondubbed films were played for 3 minutes and 15 sec with 5-sec presentations of the soundtracks from each. The observer monitored the subject's fixations throughout the procedure. Nineteen of the 20 subjects completed the search phase of the study.

Results

Table 2 displays the mean looking preferences for the sound-synchronous films for this study along with those of Experiment 1. Subjects showed no significant looking preferences for either the sound-synchronous films or the sound-appropriate films in either viewing period. The proportion of total looking time to the sound-synchronous film for viewing 1 was .559, $t(19) = 1.107$, $p > .05$, and for viewing 2 was .445, $t(19) = 1.006$, $p > .05$. Results of the search phase are expressed in terms of the proportion of first looks directed toward the previously synchronized films. These results also failed to reach significance, $t(18) = 1.157$, $p > .05$. Thus infants showed no significant looking preferences for the synchronous films when the sounds were inappropriate to the substance of the objects.

A two-way analysis of variance tested for side or film preferences during viewing period 1. No significant effects were found for side, film or interaction, $F(1,16) = .051$, $p > .05$, $F(1,16) = .484$, $p > .05$, $F(1,16) = .742$, $p > .05$, respectively. In order to determine whether infants showed any trends over time in looking to the sound-synchronous films during either viewing of the preference phase, two-sample matched t-tests were performed. The 60-sec viewing periods were divided into 30-sec segments and looking proportions for the sound-synchronous films were derived for each infant for each segment. Results indicated no trends over time for viewing period 1, $t(19) = 1.659$, $p > .05$, or for viewing period 2, $t(19) = .977$, $p > .05$.

As a test of the main hypothesis the results of this study were compared with

TABLE 2
Experiments 1 and 2: Looking Times, Preferences, and Search Proportions for the Sound-Synchronous Films

	Viewing 1 (first 60 sec)		*Viewing 2 (first 60 sec)*		*Search*[a]	
	Exp 1	*Exp 2*	*Exp 1*	*Exp 2*	*Exp 1*	*Exp 2*
Mean looking time (sec)						
Sound film	31.8	31.6	20.0	21.0		
Silent Film	20.3	24.1	27.1	28.9		
Total	52.1	55.8	47.1	49.9		
Mean preference[b]						
Proportion	.61	.56	.43	.45	.57	.54
SD	.19	.24	.26	.24	.12	.15
t	2.69*[c]	1.11	1.29	1.01	2.40*[c]	1.16

[a]These figures are based on the proportion of first looks to the sound specified films.
[b]Looking preferences were computed by dividing the number of seconds each subject looked at the sound film by the total time spent looking at both films, and averaging over the 20 values. These average proportions differ from the proportions derived from "Sound film" to "Total" above them because the mean of ratios is not, in general, equal to the ratio of means.
[c]*$p < .05$.

those of Experiment 1. Because the synchronized sounds specified objects of the wrong substance in Experiment 2, weaker preferences and attenuated search performance were predicted for Experiment 2 as compared with Experiment 1. As can be seen from Table 2, all results fell in the predicted direction. Two-sample t-tests, however, revealed that none of these differences was significant.

These data can also be analyzed for the effects of sound appropriateness when synchrony was mismatched. Were looking preferences for the sound appropriate film significantly depressed when synchrony was mismatched (Experiment 2) versus matched (Experiment 1)? In Experiment 2 the sound-appropriate film was the one that was not synchronized with the soundtrack. The looking preference for that film was obtained by subtraction, $1 - .559 = .441$, for the first viewing period. A two-sample t-test indicated that looking to the sound-appropriate film was significantly depressed in Experiment 2 where synchrony was mismatched as compared with Experiment 1 where the soundtrack was both synchronous and appropriate to the substance of the objects, $t(36) = 2.536$, $p < .02$. Results of the search phases of the two studies also differed significantly. The proportion of first looks to the sound appropriate films for Experiment 1 was .572 and for Experiment 2 was .461, $t(33) = 2.459$, $p < .02$. These results indicate that infants show significantly higher looking preferences for films whose soundtracks are both synchronous and appropriate to the substance of the objects than to films whose soundtracks are appropriate but not synchronous. They suggest that infants detect an intermodal relationship between a film and a soundtrack only when both kinds of bimodal information are consistent, and not on the basis of one kind of information alone.

Closer inspection of the looking preferences obtained in the two studies revealed an interesting pattern of results. For viewing period 1 of the preference phase the distribution of individual looking preferences seemed bimodal for subjects in Experiment 2, while those of Experiment 1 did not. Also the standard deviation of these preferences was larger for Experiment 2 (.24) than for Experiment 1 (.19). The distribution of individual looking preferences for Experiments 1 and 2 are displayed in Fig. 2. It is apparent that 11 of the 20 subjects in Experiment 1 fall near the mean, in the range of .5 to .75, whereas only 4 fall within this range for Experiment 2. A nonparametric test, the Moses test of extreme reactions (Siegel, 1956), was conducted to determine whether the two samples were drawn from populations whose distributions differed significantly, one exhibiting a more centralized distribution and the other a more dispersed one. A significant difference between the two samples was found ($p < .01$), indicating that they were indeed drawn from two different populations.

Discussion

Results of Experiments 1 and 2 taken together suggest that infants detect both synchrony and substance information. Infants looked to the sounding object when the sounds were both synchronous with the motions of the object and appropriate to its substance as in the first study. They looked longer to the synchronized object than to the nonsynchronized one when the sounds were appropriate to its substance

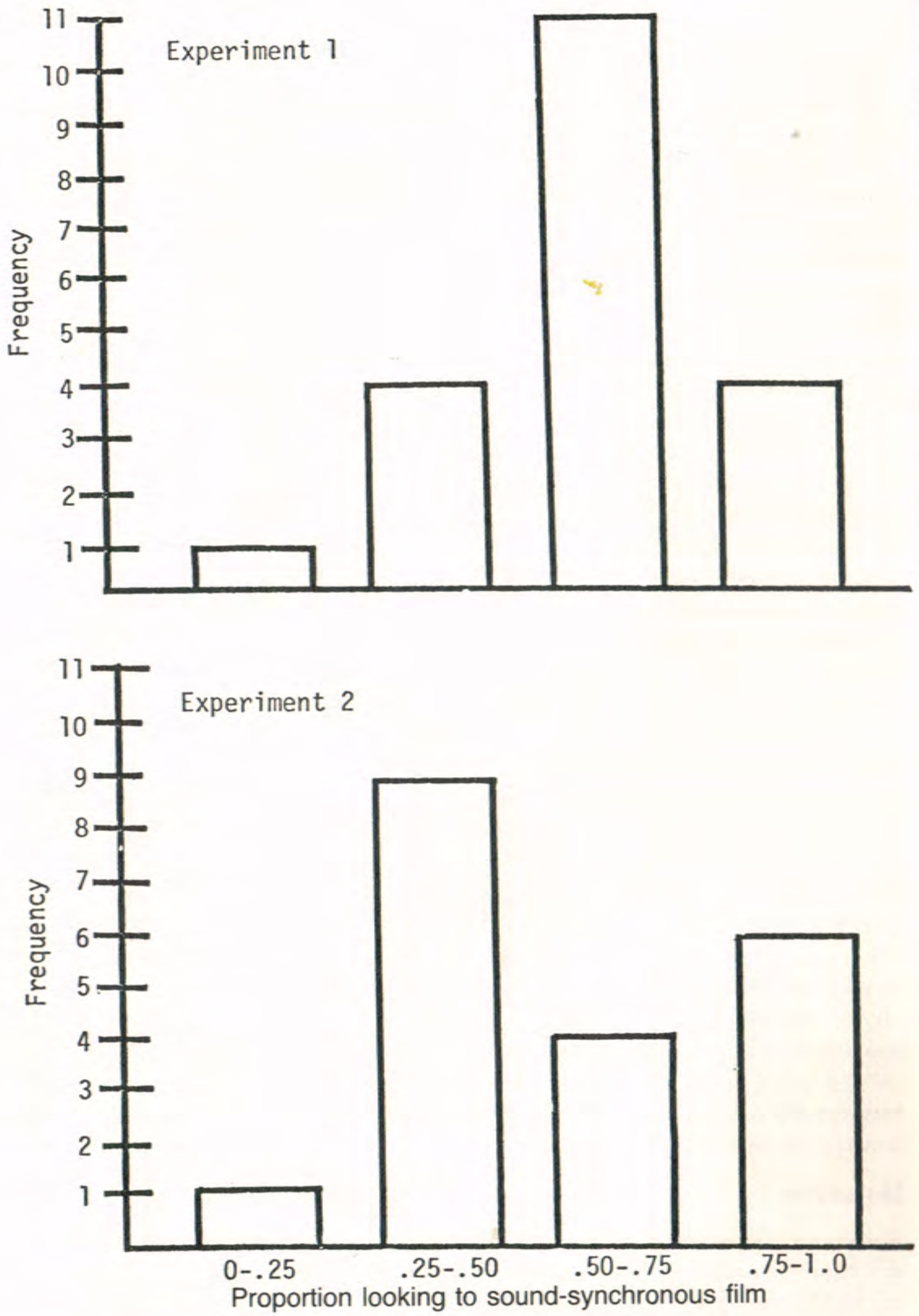

Figure 2. Experiments 1 and 2: Distribution of looking preferences for the sound-synchronous films.

(Experiment 1), but not when the sounds were inappropriate to its substance (Experiment 2). They also looked longer to the sound-appropriate object than to the sound-inappropriate one when the sounds were synchronized with the objects' motions (Experiment 1), but not when the sounds were out of synchrony with the object's motions (Experiment 2).

The results of the Moses test also suggest that the attenuated preferences in Experiment 2 were due to subjects' detection of both synchrony and substance information. Most of the subjects in this study showed extreme looking preferences, predominantly watching either the synchronous or the sound-appropriate film. Infants in this study may have been torn between a tendency to watch the sound-appropriate film versus the sound-synchronous film, with some subjects choosing one and some the other. This resulted in a bimodal distribution of looking preferences.

These interpretations, however, are not definitive. Since the soundtracks for Experiment 1 were recorded live, they were in perfect synchrony with the motions of the objects in the films. Those of Experiment 2 were dubbed, making the synchrony relationship somewhat imperfect. It is possible that infants detected this deviation and that this reduced their tendency to watch the sound-synchronous film. Although it seems unlikely that infants would be sensitive to such minor deviations in synchrony, this possibility cannot be altogether ignored. Thus, Experiments 3 and 4 address the question of infants' detection of substance and synchrony information individually.

EXPERIMENT 3

This study explored infants' detection of temporal structure specifying object substance while eliminating their opportunity for using synchrony. Recent research has shown that infants are able to differentiate between classes of rigid and deforming motions (Gibson, Owsley, Walker, & Megaw-Nyce, 1979; Walker, Owsley, Megaw-Nyce, Gibson, & Bahrick, 1980) in a visual habituation paradigm. It seemed likely that infants who were able to detect information specifying rigidity and elasticity in unimodal stimulation might be sensitive to this information in bimodal stimulation as well. The present study was designed to assess this possibility. It differed from the other studies in that the films and soundtracks were presented successively rather than simultaneously. This way, no conflicting synchrony information was present as in Experiment 2, and the effects of sound appropriateness could be assessed independent of the effects of synchrony.

Method

Ten males and 10 females participated. Their mean age was 142 days with a range of 127 to 158 days. Three subjects were exlcuded from the study due to excessive crying.

The apparatus and stimulus materials were the same as those used in Experiment 1. Each subject was first given a familiarization period with the two films. The blocks and sponges were presented silently side by side for 20 sec to acquaint the subjects with the materials and their lateral positions which were maintained throughout the testing period.

Next, infants participated in the visual preference phase. It was modified in the following ways: Subjects were given 6 trials. Each trial consisted of a 10-sec presentation of one soundtrack (from the centrally placed speaker) followed immediately by a 20-sec presentation of the two silent films side by side. The intertrial interval was approximately 5 sec long. No films or soundtracks were presented during this period. Infants received 3 trials in succession with one of the soundtracks, followed by 3 trials with the other. This method of presenting the soundtracks and films successively eliminated temporal synchrony information normally available in such events since infants never saw the films together with the sounds. In addition, the segments of a soundtrack and corresponding silent film were presented from successive portions of the film reel. This precaution ensured that no temporal pattern information was shared by a soundtrack and the corresponding portion of the silent film in each trial.

The lateral position of the two films and the order of soundtrack presentation were counterbalanced across subjects. Ten infants received the soundtrack to the blocks film first (trials 1–3) and the sponges second (trials 4–6), and 10 received the reverse order. Half the subjects in each of these groups saw the sponges on the right side and the blocks on the left, while the other half received the opposite arrangement.

The preference phase was again followed by a search phase identical to that of Experiment 1. Here, for the first time, infants were able to briefly hear the soundtracks while viewing the films. Only 15 of the 20 subjects were able to complete the

TABLE 3

Experiment 3: Looking Times, Preferences, First Looks, and Search Proportions for the Film Whose Soundtrack had Just Been Presented

	Trials 1–3	*Trials 4–6*	*Overall*	*First Looks*	*Search*[a]
Mean looking time (sec)					
Sound film	27.8	25.7	53.5		
Silent film	22.2	20.8	43.0		
Total	50.0	46.5	96.5		
Mean preference[b]					
Proportion	.57	.55	.56	.62	.62
SD	.15	.20	.12	.23	.16
t	2.08	1.01	2.20*[c]	2.27*[c]	2.89**[c]

[a]These figures are based on the proportion of first looks to the sound-specified films.
[b]Looking preferences were computed by dividing the number of seconds each subject looked at the sound film by the total time spent looking at both films, and averaging over the 20 values. These average proportions differ from the proportions derived from "Sound film" to "Total" above them because the mean of ratios is not, in general, equal to the ratio of means.
[c]$^{*}p < .05$; $^{**}p < .02$.

search phase. Five were excluded due to excessive crying. The observer monitored the infant's fixations throughout the procedure whenever the films were projected.

In sum, this experiment consisted of 3 parts; a silent 20-sec familiarization period, a modified preference phase which was made up of 6 trials of a soundtrack followed by two silent films, and a search phase like those of the other studies.

Results

Mean looking times for the modified preference and the search phases of the study are displayed in Table 3. The mean proportion of looking time to the film whose soundtrack had just been presented was .558. This proportion was significantly different from the chance value of .5, $t(19) = 2.197$, $p < .05$. Fourteen of the 20 subjects showed looking preferences of greater than .5. A second measure, the proportion of first looks (out of the 6 possible trials) directed toward the sound-specified films averaged .617. This was also significantly different from chance, $t(19) = 2.269$, $p < .05$. Subjects looked first and longer on the average to the film whose soundtrack they had just heard.

Results of the search phase were also significant. Subjects looked first toward the sound-specified films on .622 of the trials, $t(14) = 2.89$, $p < .02$. This proportion was based on only 15 of the 20 subjects, since 5 were unable to complete the search phase.

A one-way analysis of variance with repeated measures was performed to determine whether there were any significant trends over trials in looking to the sound-specified films. This analysis revealed no differences among the 6 trials, $F(5,95) = .482$, $p > .05$. Two-way analyses of variance were performed to determine whether subjects showed any side or film preferences. Separate analyses were performed for trials 1–3 and 4–6. No significant side or film preferences for trials 1–3, $F(1,16) = .309$, $p > .05$, $F(1,16) = .011$, $p > .05$, or trials 4–6, $F(1,16) = .116$, $p > .05$, $F(1,16) = .249$, $p > .05$, were found. The data from the familiarization period were also examined to determine whether subjects initially showed a film preference. A two-sample matched t-test on the raw scores for the blocks and sponges revealed no significant preferences, $t(19) = 1.256$, $p > .05$.

Discussion

Subjects looked first and longer to the sound-specified films during the preference phase and searched appropriately during the search phase. Infants were apparently able to detect information common to the types of sounds and kinds of motions the objects made; temporal information specifying their substance. This information was detected over successive presentations of the films and their soundtracks.

This audiovisual information specifying substance may be characterized as a temporal invariant. A rigid object striking a surface changes its visual trajectory rapidly and abruptly. The sound that accompanies it is abrupt and sharp and has a correspondingly rapid time course. An elastic object meeting a surface changes its visual trajectory and shape more gradually, and is accompanied by a softer more prolonged sound with a correspondingly gradual time course. This kind of temporal structure is invariant over modalities and aspects of it may specify object substance.

Little research however, has been conducted delineating the visual and acoustic information for rigid versus elastic collisions. Temporal invariants specifying object substance may be divided into at least two major categories: (a) duration: both the duration of the sound and of the visually given impact are longer for the elastic as compared with the rigid object, and (b) rate of change: both the rate of change from silence to maximum sound (rise time) and from one trajectory of motion to the other are more gradual and continuous for the elastic as compared with the rigid object. The transitions between silence and sound, and one trajectory of motion to the other are invariant over modalities. A rapid and accelerating transition may specify rigidity and a more gradual one may specify elasticity. These are but a few examples of the kinds of audiovisual invariants that may serve to specify object substance.

This experiment demonstrates that infants are able to differentiate between collisions of rigid and elastic objects on the basis of temporal structure common to vision and audition. Infants could not have detected a relationship between the films and their soundtracks on the basis of other previously defined invariants; none of them was available. The successive method of presentation eliminated temporal synchrony since the soundtracks were never played while the films were visible. Tempo of action did not differentiate the two events and the temporal pattern of the soundtrack and corresponding film on each trial did not match. Infants must have detected the bimodal temporal structure specifying object rigidity and elasticity. They picked up the temporal structure during the acoustic portion of each trial, and this guided their looking to the object of the corresponding substance during the visual portion of each trial. These results along with those of Experiment 2 demonstrate that synchrony is neither a sufficient nor a necessary condition for the infant to perceive an intermodal relationship between a sound and an object. Rather, this research suggests that infants perceive properties of objects through the detection of temporal structure. It cannot, however, be concluded from this research which of several possible temporal invariants infants detected. Infants may have simply responded to the duration of sights and sounds, their rate of change, or a more global invariant subsuming both dimensions.

The detection of this kind of temporal structure specifying object substance seems fundamental to the perception of object properties like rigidity and elasticity. Although this research along with that of Gibson et al. (1979) and Walker et al. (1980) suggests that infants perceive substantial properties of objects, this cannot be definitively concluded. One cannot demonstrate from research showing discrimination of temporal structure that object properties of rigidity and elasticity were actually perceived. Infants must have picked up the audiovisual temporal relations that specified these properties. Did infants viewing the films actually experience the rigid sounds as "hard" and the elastic ones as "soft," and did they expect the rigid object to feel hard to the touch and the elastic one to feel soft? These are intriguing questions that can only be answered by further research. Perhaps studies measuring more meaningful responses like facial expressions, surprise reactions, or approach/avoidance behavior in response to substance variation might shed more light on this question. It appears that infants by 4 1/2 months possess the underlying abilities of invariant detection and refined temporal resolution that may form the basis for the perception of meaningful objects of particular substances.

EXPERIMENT 4

Although previous research (Spelke, 1979) has demonstrated infants' sensitivity to temporal synchrony information (the coincidence between sights and sounds of impact), it is not certain from Experiment 2 whether infants can detect this information in the present more complex and unpredictable events. Experiment 4 was designed to determine this while controlling for substance information. The method was identical to that of Experiment 1. Only the film pairings differed. Two versions of the *same* event were projected side by side along with a synchronized soundtrack to one of them. It was expected that infants would demonstrate a sensitivity to synchrony information by preferentially fixating the sound synchronous films.

Method

Ten males and 10 females participated. Their mean age was 135.8 days, with a range of 127 to 149 days. Three additional subjects were excluded from the study because of equipment failure or excessive crying.

The apparatus and procedure were the same as that of Experiment 1. Only the presentation of the films differed. During the preference phase, two films of the same event were projected side by side for 60 sec along with one synchronized and appropriate soundtrack. Then two versions of the other event were shown for 60 sec along with one of their synchronized soundtracks. Thus, each infant saw two block films followed by two sponge films or vice versa. One of each pair was in phase with the appropriate soundtrack and the other was out of phase with it.

Since the temporal pattern of these events was erratic, the two projected versions of an event varied unsystematically in their phase relation to one another. Occasionally there was an accidental synchrony between the two films. For example, blocks in both versions of the blocks films might strike a surface at the same moment. Despite this occasional accidental synchrony, adults viewing the films were immediately able to determine which was synchronized with the soundtrack.

Lateral position of the sound-synchronous film and order of soundtrack presentation were counterbalanced across subjects. Half of the subjects viewed two versions of the sponges films first and the blocks second. The other half received the reverse order. Half the subjects in each of these groups heard the soundtrack synchronized with the film on the right during the first viewing period and the film on the left during the second viewing period. The other half received the opposite arrangement.

A search phase identical to that of the other experiments followed. That is, infants viewed the blocks and sponges side by side along with intermittent presentations of the soundtracks from each. The lateral positions of the sound films remained constant for a given infant throughout the preference and search phases. Eighteen of the 20 subjects completed the search phase. Two subjects were excluded, one because of equipment failure and the other because too few usable trials were obtained.

Results

Results of the study are displayed in Table 4. Contrary to what was expected, infants showed a strong and reliable looking preference for the film that was out of synchrony with the soundtrack during the first viewing of the preference phase. The proportion of total looking time to the sound-synchronous film was .399, significantly below the chance value of .5, $t(19) = 2.376$, $p < .05$. Fourteen of the 20 subjects showed this effect. During the second viewing of the preference phase, subjects again showed no looking preference for either film.

Subjects did not search reliably in this experiment. Only .494 of the infants' first looks were directed toward the sound-synchronous and appropriate films. A *t*-test indicated this search behavior to be significantly depressed relative to that of Experiment 3, $t(23) = 2.614$, $p < .02$.

Analyses of variance revealed no significant side or film preferences in viewing period 1, $F(1,16) = .112$, $p > .05$, $F(1,16) = 1.038$, $p > .05$, respectively. Two matched-sample *t*-tests were performed to determine whether subjects showed any trends over time in looking to the synchronous films,. This was calculated as before by dividing each viewing period into 30-sec intervals. Subjects showed no significant trends over time within viewing period 1, $t(19) = .537$, $p > .05$, or within viewing period 2, $t(19) = .484$, $p > .05$.

Discussion

Contrary to what was expected, infants showed a significant preference for the films that were *out* of synchrony with the soundtrack they were hearing. Although this preference was in the direction opposite of that expected, it nevertheless provides evidence of infants' sensitivity to temporal synchrony information. Since the simultaneously projected films differed only in their temporal relation to the soundtrack, infants must have differentiated between them on that basis.

TABLE 4
Experiment 4: Looking Times, Preferences, and Search Proportions for the Sound-Synchronous Films

	Viewing 1	*Viewing 2*	*Search*[a]
Mean looking time (sec)			
Synchronous film	20.3	23.0	
Nonsynchronous film	31.7	26.8	
Total	52.0	49.8	
Mean preference[b]			
Proportion	.40	.47	.49
SD	.19	.27	.11
t	2.38*	.54	.25

[a]These figures are based on the proportion of first looks to the sound-specified films.
[b]Looking preferences were computed by dividing the number of seconds each subject looked at the sound film by the total time spent looking at both films, and averaging over the 20 values. These average proportions differ from the proportions derived from "Sound film" to "Total" above them because the mean of ratios is not, in general, equal to the ratio of means.
*$p < .05$.

Why might infants have preferred to watch the nonsynchronous films? One reason may stem from the additional unpredictability or incongruity they provided. In this study the objects were of a substance appropriate to the nature of the sounds, yet their temporal pattern was discrepant. Since the events had erratic rhythms, the motions of the nonsynchronous objects were unsystematically related to the audible impact sounds; occasionally there was an accidental synchrony. This irregularity provided an intermittent intermodal relationship between the soundtrack and the nonsynchronous film. In addition, infants may have simply been bored by the two simultaneous presentations of the same film and sought more novelty. This, coupled with the intermittent intermodal relation provided by the nonsynchronous film might have caused the looking preference for nonsynchrony. A more viable explanation of these results, however, must await further research concerning the nature of infants' visual exploration: why they look at what they do.

Infants failure to search appropriately in this study was also somewhat puzzling. Only in this study did infants have the opportunity to learn which kind of object belonged with the soundtrack they were hearing during the preference phase on the basis of co-occurrence alone. Only films of the correct objects were presented along with their soundtracks. Whereas in the other studies, films of both kinds of objects were always presented along with each soundtrack. This experience with the correct object-sound pairing however resulted in no learning transfer to the search phase. This search was significantly depressed compared with that of Experiment 3 where the soundtracks and films had been presented successively in the preference phase. This suggests that infants do not perceive object-sound correspondence on the basis of contiguity alone. Infants showed no evidence of "associating" the film with its soundtrack even though they were presented together. Rather, the infant must detect audiovisual invariants specifying a relationship between a film and its soundtrack. This information may later be used to search appropriately.

One major difference between the present study and Experiments 1 and 3 could account for the discrepant search behavior. Infants in the earlier studies always viewed two different films along with a single soundtrack. This necessitated a direct visual comparison between the two films and facilitated differentiation of the objects and their locations. Successful search behavior must involve some knowledge of what each event looks like and where it is located. In the present study infants saw two versions of the same object side by side. Both objects occurred once in each location. Infants needed only to detect a synchrony relation between the object's sounds and motions. Detection of temporal synchrony does not require any attention to visual characteristics of the objects; only to their motions relative to the soundtrack.

GENERAL DISCUSSION

Several conclusions emerge from this research. To summarize, Experiment 1 tested infants' detection of audiovisual correspondence between films of rigid and elastic objects and their soundtracks. Infants showed a significant looking preference for the sound-synchronous and appropriate films. Experiments 2, 3, and 4 were conducted to further specify the nature of the audiovisual invariants detected by infants. Did subjects detect bimodal synchrony information, substance information, or both? Experiment 2 placed synchrony and substance information in conflict by playing the soundtracks in synchrony with the wrong kinds of objects. Subjects no longer showed significant looking preferences to the synchronized films (as they had in Experiment 1) when the new inappropriate sounds were played. Their looking preferences were bimodally distributed. These findings suggested that infants perceived an intermodal relationship between the objects and sounds on the basis of both synchrony and substance information. Experiments 3 and 4 were conducted to confirm this. Experiment 3 tested infants' detection of substance information in the absence of synchrony. Successive trials of a soundtrack followed by silent films of the two kinds of objects were presented. Subjects watched the film whose soundtrack they had previously heard, demonstrating detection of temporal structure specifying object substance. Experiment 4 tested infants' sensitivity to synchrony while controlling for substance information. Subjects viewed two films of the same kind of object while only one of them was synchronized with the soundtrack. They demonstrated a sensitivity to synchrony by consistently watching the film that was out of phase with the soundtrack. Infants by 4 1/2 months can detect invariant audiovisual temporal structure specifying object rigidity and elasticity. They are also sensitive to temporal synchrony in complex, unpredictable events. These findings contribute to our knowledge of the kinds of audiovisual invariants young infants are able to detect.

What implications does this early sensitivity to audiovisual invariants have for the course of intermodal development? The traditional view of this developmental course has been the "integration position" (Birch & Lefford, 1963; Bryant, 1974), which has its roots in British empiricism. According to this view, raw sensory impressions picked up through separate sensory modalities must be progressively put together through association during the course of development. Not until sensations from separate sensory channels are integrated can the child have a unified perception of a multimodal event. This view takes as its starting point for perceptual experience separate sensory channels which transmit raw sensory impressions. Research has been concerned with how and when stimulation from two modalities specifying a single property of an object comes to be integrated and treated as equivalent. A host of studies has already shown the infant to have intersensory functioning during the first months of life. To account for these findings, integration theorists must at least postulate a process of integration that develops as early as the first weeks of life.

The contrasting view of perception was advanced by Gibson (1966). This theory holds that there is direct pick up of information specifying invariant properties of objects. No processes of mediation, integration, or association are necessary for perception. He argued that although it is possible for adults to attend to sensations and sensory impressions, this is not *what* we perceive and this is not how the world is organized. Our world consists of objects and events that are perceived through unified perceptual systems. In studying the development and basis of perception we should take as a starting point, the world of multimodal objects and events and the organization inherent in it. This world presents a continuous flux of changing stimulation. The stimulation however, is highly structured; we do not have to structure it. Perceptual development proceeds through the differentiation of finer and finer aspects of the available structure (Gibson, 1969). Consequently, there is increasingly greater correspondence between perception and stimulation through the course of development. Intersensory coordination develops through the detection of invariant relations, intermodal coordinations that are present in the events themselves and specify properties of objects. As development proceeds, more invariants are differentiated. The present research was designed to demonstrate infants' sensitivity to audiovisual invariants. According to this view, if infants detect these invariant relations, there is no need to postulate a process of integration or association since they are already perceiving the multimodal object as a whole. The task is to delineate and specify the kinds of invariant relations we detect.

How can the early detection of invariants enable the infant to veridically differentiate the world of multimodal objects and events? How might infants come to perceive certain patterns of light and sound as unitary and others as unrelated? Sights and sounds are united by a common rhythmic structure, temporal synchrony, and tempo of action. Because the same structure is present in both light and sound, it sets the event apart as unitary and separates it from other ongoing events. By picking up this invariant structure, the infant is necessarily attending to the multimodal event as a whole and differentiating it from other ongoing activity.

This phenomenon was illustrated in a recent selective looking study (Bahrick, Walker, & Neisser, 1981). In order to direct infants' selectivity to one of two superimposed visual episodes, we capitalized on their tendency to abstract audiovisual invariants. A synchronous and appropriate soundtrack was played to one of the two overlapping films. This enabled the infants to successfully attend to one superimposed event while ignoring the other. The "selective looking" phenomenon was quite compelling for the adult observer. Without sound, the superimposed display appeared as several ghostly images passing through one another. The addition of the soundtrack caused the sound event to phenomenologically "leap out" from its background of the other event. Switching the soundtrack caused the other event to stand out. Adults found it difficult to attend to the display as a whole when a soundtrack was played. I imagine the experience of the infant to be something like that of the adult viewing the overlapping films when a soundtrack is played.

The addition of the soundtrack provided invariant temporal structure that united the soundtrack with the corresponding part of the visual display. Pick up of such invariant relations directs perceptual selectivity and should entail perception of unitary events from the beginning. Nothing need be progressively integrated across modalities over time. The multimodal information is already united in the structure detected over time. By differentiating unitary events in this way, the infant can begin to differentiate finer aspects of the event. Aspects that make it unique and distinct from others that he may later perceive.

The task of developmental research from an invariant detection view is to determine the nature and course of the differentiation process. What kinds of relations are differentiated first, and how does this lead to the detection of other, more specific aspects of the available stimulation? How can this process culminate in the intermodal knowledge of the adult perceiver such that eventually stimulation to one modality gives rise to specific expectations about what is to be perceived through another modality? Developmental research systematically investigating differentiation of invariant relations, and how this sensitivity emerges and is refined over several successive ages, may begin to clarify this developmental course.

Taken together with other recent findings, this research suggests a general framework for viewing the development of intermodal knowledge in infancy. The neonate explores his environment through spatially coordinated systems of vision and audition (Muir & Field, 1979; Wertheimer, 1961). Spatial properties of acoustic stimulation influence visual scanning (Mendelson & Haith, 1976). This coordination increases the probability that the infant will look to what he is hearing. The infant seeks structure in stimulation and the detection of this structure guides further exploration. Most audiovisual events can be characterized by a temporal structure that is invariant over their modality. Infants detect at least four kinds of temporal structure; temporal synchrony, tempo of action (Spelke, 1979), temporal pattern (Allen et al., 1977), and substance information (rate of change). The detection of these invariants unites the light and sound from a single event and separates it from other ongoing events that do not share this structure. Selective tuning to these temporal invariants at an early age provides an economical means of verdically differentiating events. This may in turn facilitate detection of other embedded audiovisual relations and promote differentiation of distinctive object features. A sensitivity to invariant relations may guide infants' exploration of everyday events, providing the basis for the development of the intermodal knowledge possessed by adults.

REFERENCES

Allen, T. W., Walker, K., Symonds, L., & Marcell, M. Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology*, 1977, *13*, 225–229.

Bahrick, L. E., & Walker, A. S. Infants' perception of amodal information in natural events. Paper presented at the Eastern Psychological Association meetings, Washington DC, March, 1978.

Bahrick, L. E., Walker, A. S., & Neisser, U. Selective looking by infants. *Cognitive Psychology*, 1981, *13*, 377–390.

Birch, H. G., & Lefford, A. Intersensory development in children. *Monographs of the Society for Research in Child Development*, 1963, *28*, Serial #89.

Bryant, P. *Perception and understanding in young children: An experimental approach*. London: Methuen, 1974.

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

Gibson, E. J., Owsley, C. J., Walker, A. S., & Megaw-Nyce, J. Development of the perception of invariants: Substance and shape. *Perception*, 1979, *8*, 609–619.

Gibson, J. J. *The senses considered as perceptual systems*. Boston: Houghton-Mifflin, 1966.

Mendelson, M. J., & Haith, M. M. The relation between audition and vision in the human newborn. *Monographs of the Society for Research in Child Development*, 1976, *41*, Serial #167.

Muir, D., & Field, J. Newborn infants orient to sounds. *Child Development*, 1979, *50*, 431–436.

Ryan, T. A., Joiner, B. L., & Ryan, B. F. *Minitab student handbook*. North Scituate, MA: Duxbury Press, 1976.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Spelke, E. S. Infants' intermodal perception of events. *Cognitive Psychology*, 1976, *8*, 553–560.

Spelke, E. S. Intermodal exploration by four-month-old infants: Perception and knowledge of auditory-visual events. Unpublished doctoral dissertation, Cornell University, 1978.

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

Walker, A. S., Owsley, C. J., Megaw-Nyce, J., Gibson, E. J., & Bahrick, L. E. Detection of elasticity as an invariant property of objects by young infants. *Perception*, 1980, *9*, 713–718.

Watson, J. S. Operant fixation in visual preference behavior of infants. *Psychonomic Science*, 1968, *12*, 241–242.

Wertheimer, M. Psychomotor coordination of auditory-visual space at birth. *Science*, 1961, *134*, 1692.

29 April 1982; Revised 29 September 1982 ■