Developmental Psychology
1979, Vol. 15, No. 6, 626-636

Perceiving Bimodally Specified Events in Infancy

Elizabeth S. Spelke
University of Pennsylvania

Four-month-old infants can perceive bimodally specified events. They respond to relationships between the optic and acoustic stimulation that carries information about an object. Infants can do this by detecting the temporal synchrony of an object's sounds and its optically specified impacts. They are sensitive both to the common tempo and to the simultaneity of such sounds and visible impacts. These findings support the view that intermodal perception depends at least in part on the detection of invariant relationships in patterns of light and sound.

Humans live in a world of objects and events that can be seen, heard, and felt. When mature perceivers look and listen to an event simultaneously, they experience a unitary episode. When they look at one event while listening to another, they are aware of two separate happenings. These experiences are possible because adults can determine if simultaneous patterns of light and sound are produced by a single object. Adults can perceive bimodally specified events.

What are the origins of this capacity? Many philosophers and psychologists have suggested that it arises from experience. Perceivers come to relate visual and auditory sensations through direct association (Berkeley, 1709/1910; Birch & Lefford, 1967; Mill, 1829), verbal mediation (Blank & Bridger, 1964), or the integration of schemes for looking and listening (Piaget, 1952). According to any of these views, humans begin life experiencing unrelated sensations in each sensory modality. They must learn to put together the separate experiences provided by each sense.

The present research explored a contrasting position, developed by James J. and Eleanor J. Gibson. This position derives from the theory that perceiving depends on the detection of invariants in stimulation (E. Gibson, 1969, E. Gibson, Note 1; J. Gibson, 1966, 1979). To the Gibsons, a stimulus invariant is a higher order relationship that remains constant as other stimulus variables change. Information is invariant over the auditory and visual modalities when the same relationship characterizes stimulation both to the eye and to the ear. The Gibsons assert that perception of bimodally specified events depends on detection of such invariants. If a person discovers the same relationship by looking and listening, he or she will perceive a unitary event. If no such relationship is detected, he or she will perceive two unrelated events, one specified in light and the other in sound. Even newborn infants, according to this view, will perceive bimodally specified events if they are sensitive to the appropriate stimulus invariants.

To be interesting and testable, any theory of intersensory development must be specific. A theory of association by contiguity must specify how auditory and visual arrays are parsed into elements that can be associated. A theory of scheme integration must

Portions of this research are based on a thesis submitted in partial fulfillment of the doctoral degree at Cornell University. Portions appeared in a paper presented at the International Conference on Infant Studies, Providence, Rhode Island, March 1978. The research was supported by a grant to Eleanor J. Gibson from the National Science Foundation (76-14-942), by a predoctoral fellowship from the National Institute of Mental Health (MH-05707), and by biomedical research funds from the University of Pennsylvania.

I am grateful to Ulric Neisser and to Eleanor J. Gibson for guidance, and to Eleanor J. Gibson and James J. Gibson for criticisms of an earlier draft of this article.

Requests for reprints should be sent to Elizabeth S. Spelke, Department of Psychology, University of Pennsylvania, 3813-15 Walnut Street T3, Philadelphia, Pennsylvania 19104.

Copyright 1979 by the American Psychological Association, Inc. 0012-1649/79/1506-0626$00.75

describe the detailed characteristics of the young child's schemes for looking and listening, and it must indicate which aspects of each activity become reciprocally assimilated. An invariant detection theory must specify the class of stimulus relationships that can serve as invariants. Without such constraints, each of these theories could be stretched to account for the results of virtually any developmental experiment. In this presentation, one specific version of invariant-detection theory is proposed. Tests of one of its predictions are then described.

The present proposal was derived from the attempt by J. Gibson (1966, 1979) to give limits and substance to invariant-detection theory through an ecological analysis of visual perception. As perceivers evolved, Gibson suggests, they became sensitive to the optic invariants that specified the significant properties of their visual environment. Experimental studies of "ecological optics" take as invariants those stimulus relationships that are specific to the objects, events, and surface layout that an animal must perceive (cf. Fieandt & Gibson, 1959; Gibson, Owsley, & Johnston, 1978).

Although Gibson has not extended this ecological analysis to the perception of bimodally specified events, such an extension seems feasible. The human capacity for detecting invariants over patterns of light and sound might be constrained by two characteristics of natural objects. First, sounds generally emanate from the visible direction of the object that they specify. Optic and acoustic stimulation specifying an object are localized in a single position in space, and their directions change in concert as the object or the observer moves. Second, most sounds are produced by moving surfaces that cause vibrations in the air. These motions, when visible, are temporally synchronized with the sounds that specify them. Perceivers might therefore be sensitive to *spatial invariants*, uniting the audible and visible direction of an object, and to *temporal invariants*, uniting its audible and visible motions. Humans might innately perceive a bimodally specified event when one or both kinds of intermodal relationship are detectable. The present research investigated whether young infants can perceive an auditory–visual relationship by detecting a temporal invariance.

These experiments were based on several recent investigations of auditory–visual perception in infancy. Studies of exploration provide evidence that looking and listening are coordinated early in life. Young infants increase their visual activity at the time that a sound occurs (Haith, 1973; Horowitz, 1974). They also tend to look in the spatial direction of a sound under some circumstances (Butterworth & Castillo, 1976; Field, DiFranco, Dodwell, & Muir, 1979; Mendelson & Haith, 1976; Muir & Field, 1979; Wertheimer, 1961). A number of investigators have used infants' proclivity for auditory–visual exploration to probe their perception of bimodally specified events (Lyons-Ruth, 1977; Spelke, 1976; Lewis & Hurowitz, Note 2).

Spelke (1976) gave 4-month-old infants a test of visual preference between films of a game of peekaboo and of a music sequence played on toy percussion instruments. As the films were projected side by side, the sound track of one event was played through a centrally placed speaker. The infants looked toward the acoustically specified event for a longer time. They were able to determine, on some basis, which of the two filmed episodes corresponded to each sound accompaniment. A subsequent experiment indicated that infants responded only to the intermodal relationship in the percussion sequence (Spelke, 1978). In further research, infants detected relationships between optic and acoustic stimulation specifying other natural events like the percussion episode. Bahrick, Walker, and Neisser (Note 3) presented 4-month-old infants with all pairings of three filmed events: a game of pat-a-cake, a slinky toy repeatedly opened and closed by two hands, and a musical sequence played on a xylophone. Each pair of events was presented with one of the two appropriate sounds. Infants reliably preferred the acoustically specified event in five of the six experimental conditions. Young perceivers are evidently able to detect auditory–visual relationships in a variety of episodes.

These studies indicate that 4-month-old infants can determine that an acoustic pattern is related to one optically specified event and not to another. The experiments did not reveal how infants discovered the intermodal relationships. Because the events were natural, and possibly familiar, infants might have learned in the past about the relevant auditory–visual correspondences. A 4-month-old infant might have learned, for example, that clapping noises are produced by objects looking like hands. An alternative explanation does not depend on the assumption of specific prior experience with these events. In each episode, sounds were synchronized with the visible motions of objects. Infants might have discovered the auditory–visual relationships by detecting this temporal invariance. The research described herein focused on the latter possibility.

Three experiments investigated infants' capacity to perceive bimodally specified events by detecting the temporal synchrony of sound bursts with the visible impacts of surfaces. Four-month-old infants were presented with two objects and two sounds that were paired in unfamiliar combinations. Each object bounced against a surface and made a different percussion sound. One sound was played while both objects were visible and was centered between them. The sounds and visible impacts were temporally related in different ways in each of the experiments. In Experiment 1, each sound occurred at the same tempo as the impacts of one object and was simultaneous with those impacts. Sounds and impacts were not simultaneous in Experiment 2, but they continued to be united by their common tempo. Sounds and impacts were again simultaneous in Experiment 3, but they shared no distinctive tempo. No further information united the optic and acoustic stimulation in any study. The quality of the sound was such that it could have been produced equally well by the bounce of either object. Infants could respond to the auditory–visual relationships only if they detected the temporal invariants.

If infants are found to be insensitive to this temporally invariant information, then we may reject one version of the invariant-detection hypothesis. The earliest discovery of auditory–visual relationships in events like pat-a-cake will be shown not to depend on the detection of the synchrony of sound and movement. If infants are found to be sensitive to temporal invariants, then the invariant-detection approach to intermodal perception will remain a viable position. Further studies could then attempt to settle a more crucial question for theories of intersensory development: Is invariant detection a primary ability that serves as a basis for the discovery of all auditory–visual relationships? Or is it a secondary ability acquired by infants who have learned to relate what they see and hear in some other way?

Experiment 1

Infants viewed films of two objects bouncing at different rates and producing synchronized sounds. During a *preference episode* and a *search episode*, the babies were presented with both filmed events, side by side, while they heard the sound track for each event in turn through a centrally placed speaker. Visual attention to the events was monitored. An infant who detected the synchrony of sound bursts and visible impacts was expected to look toward an event when its synchronized sound was played.

Method

Subjects. Sixteen infants aged 3 months 29 days to 4 months 12 days (mean age, 4 months 5 days) participated in the experiment. One additional baby was excluded from the sample because of experimenter error. All infants were full-term, with no defects in vision or hearing reported by the parents or apparent to the experimenters. The infants resided in or near Ithaca, New York. No attempt was made to balance or control for their race, sex, or socioeconomic status.

Display materials. The infants were presented with color motion-picture films. In each film, a toy stuffed animal—a yellow kangaroo or a gray donkey—appeared on a grassy lawn. The animal was lifted into the air and dropped to the ground, via thin puppet strings, by an assistant standing offscreen. Each animal was moved at a regular rate of one bounce per 2 sec (slow tempo) or two bounces per sec (rapid tempo). Each impact with the ground was accompanied by a burst of sound: a "thump" for one animal and a "gong" for the other. The thump was actually produced by hitting a shoe against a hollow wooden box; the gong occurred when the shoe was hit against the lid of a metal oil drum. Each auditory accompaniment was recorded on the sound track of the appropriate film. Four films

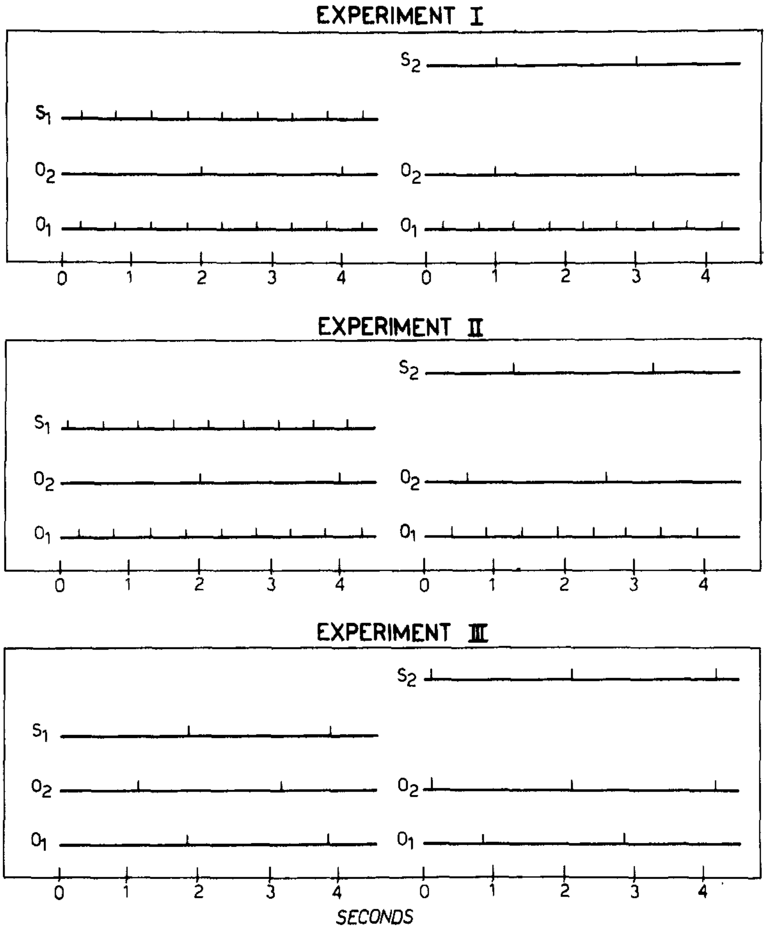

Figure 1. Schematic representation of the temporal relationships in (a) Experiment 1, (b) Experiment 2, and (c) Experiment 3. (O_1 and O_2 are the two objects, and S_1 and S_2 are the corresponding sounds.)

were made in all, one of each animal moving at each tempo. The thump accompanied the rapid tempo events, and the gong accompanied the slow tempo events. Each infant viewed two of the four films. He or she saw two different toy animals moving at two different rates and accompanied by two different sounds. The temporal relationships in these events are depicted in Figure 1a.

The films were rear projected onto the left and right halves of a translucent 80 cm × 50 cm screen by means of sound projectors. The filmed images each measured 36 cm × 33 cm and were 8 cm apart. A flashlight mounted between the images could be flickered to attract the baby's attention. The sound track of either film was played through a speaker placed 1.5 m behind the center of the screen at a volume averaging 66 dB (A) at the infant's location. The ambient noise level in the room averaged 42 dB when silent films were projected. A baby was seated in a reclinable infant seat with his or her head about 40 cm from the center of the screen. Time of looking toward each filmed event was recorded by experimental assistants who observed the infant through peepholes below the screen. They depressed buttons connected to an event recorder on which the onset of each sound track was also recorded.

When the films were shown to an infant, small and irregular variations in the speeds of the projectors caused the tempo of the events to vary slightly. No effort was made to synchronize the onsets of the two films or the speeds of the projectors. The phase relations between the events, therefore, varied across subjects and changed unsystematically over the course of each viewing period.

Design and procedure. Each infant participated in a visual-preference episode followed by a visual-search episode. The preference episode used the method of Spelke (1976). Each infant was presented with two events, side by side, while the synchronized sound of one event was played through the speaker for 100 sec. After a pause of 10 sec (or longer if a baby was fussy), the films were projected again for 100 sec with the sound track to the other film.

The visual search episode followed a 5-min intermission. It was adapted from the experimental method of

Spelke and Owsley (1979). The same two films were projected silently, and the light between them was flashed for 1 sec. The sound track to one film was then played for 5 sec, beginning when the flashing ceased. When the sound ended, the light returned for 1 sec, and a second search trial—5 sec of one of the two sound tracks—followed. Trials were given repeatedly throughout the episode, which lasted 200 sec. The two filmed events were continuously visible during this period. Between 8 and 12 search trials were given with each sound track, the sounds occurring in a different order for each infant. The orders of sound tracks were random, with the restrictions that neither was played more than three times in succession and that each was played an equal number of times.

Each infant viewed the same pair of films in the same lateral positions throughout the experiment. The pairing of sounds and objects, the lateral position of the first sound-related film, and the sound order during the preference episode were counterbalanced across infants. Different observers recorded looking during the preference and search episodes. Neither observer was aware of the lateral position of each acoustically synchronized event, and the observer for the search episode was also unaware of the baby's performance during the preference episode. Five observers assisted in the experiment. Reliabilities of each observer with at least one other observer were calculated for two or more experimental sessions. A reliability was expressed as the proportion of seconds during which the observers agreed on the direction of the infant's looking. For the preference episode, the reliabilities between pairs of observers ranged from .90 to .99 and averaged .94; reliabilities on the search episode ranged from .93 to .99 and averaged .96.

Dependent measures and data analysis. For each session of the preference episode, the proportion of looking toward the acoustically specified event was calculated for each infant. This proportion was derived by dividing the infant's total amount of looking time toward either event into the amount of looking time toward the synchronized event. Mean preferences were also calculated for each infant by averaging the proportion scores for the two sessions.

A trial of the search episode was not scored if the infant was already looking toward either event when the sound track began; an average of 8.9 trials per infant remained for analysis. Four measures of visual search were derived from looking patterns on these remaining trials: (a) *first look*—the number of trials on which the infant looked first toward the acoustically synchronized event and the number of trials on which he or she looked first toward the nonsynchronized event; (b) *eventual look*—the number of trials on which the infant looked at all (first or second) toward the synchronized and nonsynchronized events; (c) *latency of looking*—the mean duration of time that elapsed between the onset of a sound and the infant's first look toward the synchronized event and, similarly, the elapsed time between sound onset and the first look toward the nonsynchronized event; and (d) *duration of looking*—the mean amount of looking time that the infant devoted to the synchronized event and to the nonsynchronized event. Only looks occurring within 5 sec after sound onset were scored. If no look toward a given event occurred within 5 sec, the infant was given a latency score of 5 sec and a duration score of 0 sec for that event on that trial. The data from the preference and search episodes were reduced by experimental assistants. They analyzed looking to the left and right on trials with the thump and gong sound tracks, without knowing which sound specified the object on the left and which sound specified the object on the right.

Visual preferences were tested against the chance value of .500 by t tests. Preference for the synchronized event was analyzed by a 2 (lateral position of the synchronized event) × 2 (animal synchronized with the sound) × 2 (sound quality and tempo) factorial analysis of variance on each session of the preference episode. Visual search was analyzed by t tests that tested the difference in looking time toward the acoustically specified and nonspecified events on each of the four search measures against the chance value of 0.

Results

Preference episode. The results of the two preference sessions appear in Table 1. Infants exhibited a visual preference for the acoustically synchronized event during the first session. During the second session, they exhibited no such preference. The mean preference for the synchronized events was marginally significant. Subsequent analyses revealed no main effects or interactions of the lateral position of the sound film, the animal in the film synchronized with the sound, or the sound rate or quality on preference for the synchronized event during the first session. The same was true for the second session, except that preference for the synchronized event was greater if the object was the kangaroo, $F(1, 8) = 7.96$, $p < .05$.

Search episode. The principal results of the search episode appear in Table 2. Infants looked first toward the event specified by each brief sound track on reliably more trials, and they looked toward that event more quickly. They also tended to look eventually toward the acoustically specified event on more trials, but this tendency was not significant. There was virtually no difference in the duration of looking toward the synchronized and nonsynchronized events. By the first-look measure, 10 infants searched on more trials for the synchronized event, 4 searched on more trials for the nonsynchronized event, and 2 searched equally often for each event.

Discussion

The infants in Experiment 1 were sensitive to a temporal invariance in the flow of optic and acoustic stimulation. They perceived a bimodally specified event when an object's sounds and impacts were synchronized. Detecting this synchrony guided their looking during the search episode. The infants tended to look for an object when its synchronized sound was briefly presented to them. The clearest measures of the infants' search for a sounding object were the first-look and latency measures. Infants looked quickly toward an object when they detected its sound, but they did not look for it more frequently or sustainedly.

The results of this experiment fail to agree, in one respect, with those of earlier research (Spelke, 1976; Bahrick et al., Note 3). Infants preferred the acoustically specified events much less strongly in the present study. Preferential looking toward the synchronized event was obtained during the first preference session, but not during the second preference session or during the search episode (on the duration-of-looking measure). Several factors could account for this difference. First, the events in the present study were more similar to each other than were the events presented to infants in the earlier experiments. Infants may look longer toward an acoustically specified event only if the alternative, nonspecified event differs markedly from it. Second, infants might have known something about the events in the earlier experiments—clapping hands, talking people, and the like—before those studies began. In contrast, infants could not have known previously about the sound–object pairings in the present study. Visual preference for an audible event may depend on such prior knowledge. Third, the kangaroo and donkey events were more repetitive than the events in the other experiments. Infants may have attempted to keep track of both of them by dividing their looking time between the acoustically specified and the nonspecified episodes.

Table 1
Visual Preference for Acoustically Specified Events: Experiment 1

Session	Looking time (sec): Specified event	Looking time (sec): Nonspecified event	Preference	$t(15)$
1	52.9	32.6	.614	1.87**
2	35.5	41.4	.476	<1
M	44.2	37.0	.545	1.57*

* $p < .10$, one-tailed. ** $p < .05$, one-tailed.

Table 2
Visual Search for Acoustically Specified Events: Experiment 1

Measure	Specified event	Nonspecified event	$t(15)$
First look (no. trials)	4.44	3.50	2.34*
Eventual look (no. trials)	5.56	4.94	1.30
Latency (sec)	2.25	2.68	1.75*
Duration (sec)	1.49	1.36	<1

* $p < .05$, one-tailed.

Despite the equivocal results of the preference episode, the search episode revealed that infants can perceive a bimodally specified event by detecting the synchrony of sound bursts and visible movements. Infants could do this in either of two ways. First, the two filmed objects moved at different rates. Infants may have detected a relationship between sounds and visible impacts that occurred in a common tempo. Habituation research has revealed that 7-month-old infants respond to an invariant rhythmic sequence of lights and tones (Allen, Walker, Symonds, & Marcell, 1977). The present findings could reflect a similar response to an invariant tempo of sound and movement. Second, each sound burst occurred whenever one of the objects contacted the ground. The burst was not related in time to the position of the other object. Infants may have detected a relationship between sound bursts and visible impacts that occurred simultaneously, irrespective of their tempo.

It is important to distinguish these two

Table 3
Visual Preference for Acoustically Specified Events: Experiment 2

Session	Looking time (sec): Specified event	Looking time (sec): Nonspecified event	Preference	$t(15)$
1	37.9	41.1	.468	<1
2	36.5	32.5	.541	<1
M	37.2	36.8	.504	<1

possibilities because the latter is consistent with one version of an associative-learning hypothesis. One might posit that infants (a) parse a visible event into impacts and times between impacts, (b) parse a stream of sound into bursts and pauses, and (c) form associations between sound bursts and impacts. Associative learning of this kind could not account for perception of an intermodal relationship if sound bursts and visible impacts occurred at different times but at a common rate. Experiment 2, accordingly, probed infants' sensitivity to the common tempo of sounds and visible impacts, using events in which sounds and impacts were not simultaneous.

Experiment 2

This study followed the method of Experiment 1 with one modification. Each sound track was played out of phase with the filmed event that it specified, so sounds and visible impacts did not occur simultaneously. Only the rate of movement united each sound track with one filmed object. Figure 1b depicts the temporal relationship in schematic form.

Method

Subjects. Sixteen infants aged 3 months 15 days–4 months 29 days (mean age, 4 months 3 days) contributed to the experiment. Two additional babies failed to complete the study because of fussiness. All infants were full-term with apparently normal vision and hearing. They lived in or near Philadelphia, Pennsylvania.

Display materials and apparatus. The filmed events of Experiment 1 were shown by means of sound projectors while tape recordings of each sound track were played through a centrally placed external speaker. The films and tape recordings were begun at haphazard locations and were not mechanically synchronized. Thus, each sound track occurred out of phase with the impacts of each object, the objects moved out of phase with each other, and all phase relations changed unsystematically over the course of a session due to variations in the speed of the projectors and tape recorder. Each sound track occurred at approximately the same rate as the impacts of one filmed object.

Although the experiment took place in a different laboratory, the apparatus was essentially the same as in Experiment 1. One change should be noted: The flashlight, which preceded the onset of each trial in the visual-search episode, was replaced by a more attractive vertical row of 14 colored Christmas lights mounted between the two filmed images.

Design, procedure, and analysis. Except as noted, the method of Experiment 1 was followed. Each sound track was played over the tape recorder during both preference and search episodes. During the search episode, an average of 10.1 usable trials were administered to each infant. Every baby was observed by two assistants who independently recorded looking toward the two events. Their reliabilities during the preference episode ranged from .63 to .96 and averaged .88; reliabilities during the search episode ranged from .74 to .95 and averaged .80.

Results

Preference episode. The results of the preference episode appear in Table 3. There was no tendency to look longer toward the object that moved at the rate of the concurrent sound bursts during either preference session. Subsequent analyses revealed that infants looked at the acoustically specified event more during the first and second sessions, if that event occurred at the rate of the rapid thump sound, $F(1, 8) = 12.89$ and 15.48, respectively, $p < .01$. This effect reflects an overall preference for the rapid-tempo event, irrespective of sound. No other factors influenced visual preferences.

Search episode. The results of the search episode appear in Table 4. Infants looked first and eventually on more trials toward the acoustically specified event than toward the nonspecified event. They tended to look toward the former event more quickly as well, but the difference in latency was not reliable. Once again, there was no effect of the auditory accompaniments on the duration of looking. Ten infants looked first on more trials toward the synchronized event, five looked first on more trials toward the nonsynchronized event, and one looked first equally often toward each event.

Since the preference episode had revealed an effect of movement tempo on infants' looking patterns, the effects of tempo were further analyzed on the two statistically reliable search measures (first looks and eventual looks). On both measures, infants looked toward the rapid-tempo event on more trials than toward the slow-tempo event irrespective of sound accompaniment, $t(15) = 2.50$ and 1.90, respectively, $p < .05$. The tendency to look toward each event was greater, however, when the appropriate sound was played. As Table 5 indicates, the effect of the slow sound track on looking toward the slow-tempo event was more reliable than the effect of the rapid sound track on looking toward the rapid-tempo event. Infants thus detected the auditory-visual relationship in the slow-tempo event. Their ability to detect this relationship in the rapid-tempo event is less clear.

Discussion

Four-month-old infants perceived bimodally specified events by detecting the common rate of sound and visible movement. They searched for the object moving at the same tempo as a sequence of sound bursts, even though the bursts were not simultaneous with that object's visible impacts. Infants evidently can detect an auditory-visual relationship when sounds and visible movements occur at the same rate.

The search by infants in Experiment 2 would be difficult to explain within the framework of traditional association theories. Infants could not have formed associations between individual sound bursts and visible impacts, since those sounds and impacts were not temporally contiguous. Each sound burst was just as likely to occur at the time of the inappropriate object's impact as at the time of the appropriate object's impact. Only the relationship between successive impacts united each sound to one visible object. Infants detected this relationship.

The first-look, eventual-look, and (to a lesser extent) latency measures of the search test reflected the infants' perception of auditory-visual relationships. As in Experiment 1, infants looked more readily, but not for a longer duration, toward the object that was specified by each sound. They did not look longer at the acoustically specified event during the preference episode or the search episode. The similarity or repetitiveness of the kangaroo and donkey sequences or the artificiality of the sound-object pairings may account for the absence of a reliable visual preference.

Table 4
Visual Search for Acoustically Specified Events: Experiment 2

Measure	Specified event	Nonspecified event	$t(15)$
First look (no. trials)	5.00	3.50	2.14**
Eventual look (no. trials)	6.75	5.75	2.33**
Latency (sec)	3.02	3.34	1.46*
Duration (sec)	1.09	1.16	<1

* $p < .10$, one-tailed. ** $p < .05$, one-tailed.

Preference and search for acoustically specified objects were affected in complex ways by an object's rate of movement. In the search episode, infants appeared to respond more reliably to the auditory-visual relationship in the slow-tempo event. This finding must be interpreted with caution, however, since the fast and slow events were not of equal interest to the subjects. An effect of sound on looking toward the rapid-tempo event may have been less detectable because of the high base rate of looking toward that event. Infants' perception of sounding objects moving at different tempos merits further investigation with events of equal intrinsic interest to the subjects.

Experiment 3

Experiment 3 used the preference and search methods to investigate sensitivity to the simultaneity of sound bursts and visible impacts. Four-month-old infants were presented with two animals moving at the same rate, each filmed with a different synchronized auditory accompaniment. Since the tempos of the two events did not differ, only the simultaneity of sounds and impacts

Table 5
Visual Search for Rapid- and Slow-Tempo Events: Experiment 2

	Rapid-tempo film			Slow-tempo film		
Measure	Rapid sound	Slow sound	$t(15)$	Slow sound	Rapid sound	$t(15)$
First look	2.87	2.12	1.63*	2.12	1.37	2.14**
Eventual look	3.62	3.19	1.11	3.12	2.56	2.75***

* $p < .10$, one-tailed. ** $p < .05$, one-tailed. *** $p < .01$, one-tailed.

united each sound track with its synchronized film. The temporal relationship is depicted schematically in Figure 1c.

Method

Subjects. Sixteen healthy infants from Philadelphia, Pennsylvania, or its suburbs participated in the experiment. One additional infant failed to complete the study because of fussiness. The ages of the infants in the final sample ranged from 3 months 23 days to 4 months 18 days and averaged 4 months 6 days.

Display materials and apparatus. The films and laboratory facilities were those of Experiment 2. Only the films of the slowly moving objects were used. Each infant viewed the kangaroo and donkey, one synchronized with the thump sound and one with the gong sound. The pairings of sounds and objects were counterbalanced across infants. As in the previous studies, the films were begun at haphazard locations and the projectors were not synchronized. Hence, the phase relationship between the two events varied from subject to subject and changed over the course of an episode.

Design, procedure, and analysis. The method followed that of Experiment 1. Films were projected with synchronized sound tracks. During the search episode, infants received an average of 8.6 usable trials. Every infant was observed by two assistants with reliabilities ranging from .80 to .98 (*M*, .91) for the preference episode and .82 to .97 (*M*, .89) for the search episode.

Results

Preference episode. Looking preferences are given in Table 6. Infants exhibited a visual preference for the acoustically synchronized event in each session, but this preference was only reliable when the data from both sessions were combined. Subsequent analyses revealed no effect of the lateral position of the synchronized event, the object depicted in that event, or the quality of the sound on infants' preferences during either session.

Search episode. Infants searched reliably for the acoustically specified event, as indicated in Table 7. They looked first and eventually toward the synchronized event on more trials, and they looked toward that event more quickly. They did not look at the synchronized event for a reliably longer duration. On the first-look measure, 11 infants searched more for the synchronized event, 3 searched more for the nonsynchronized event, and 2 searched equally for each event.

Discussion

Infants were able to detect the simultaneity of sound bursts and the visible impacts of objects even when the synchronized and nonsynchronized objects moved at the same rate. They revealed this ability most clearly in the visual-search episode. By detecting the synchrony of sounds and impacts, infants were able to look for an event when its sound was briefly played. When sound bursts and visible impacts were simultaneous, the infants perceived a bimodally specified event. The first-look, eventual-look, and latency measures provided the best indexes of visual search. As in Experiments 1 and 2, the duration measure proved not to index search at all.

Table 6
Visual Preference for Acoustically Specified Events: Experiment 3

	Looking time (sec)			
Session	Specified event	Nonspecified event	Preference	$t(15)$
1	47.0	32.7	.583	1.53*
2	54.6	33.1	.545	<1
M	50.8	32.9	.566	2.06**

* $p < .10$, one-tailed. ** $p < .05$, one-tailed.

Infants exhibited a visual preference for the acoustically synchronized events as well, but this preference was exhibited only on one of three measures and, hence, was not convincingly strong. No such preference was found in a recent replication of this experiment (Spelke, Note 4), although the search episodes of the two studies produced results that agreed closely. Visual preference for acoustically specified events was distinctly weaker in the present study than in the experiments by Spelke (1976) and by Bahrick et al. (Note 3). Despite their weak visual preferences, however, infants exhibited strong and consistent visual search for the synchronized objects.

Table 7
Visual Search for Acoustically Specified Events: Experiment 3

Measure	Specified event	Nonspecified event	$t(15)$
First look (no. trials)	4.12	3.06	1.93*
Eventual look (no. trials)	5.69	4.75	2.09*
Latency (sec)	2.85	3.14	2.41*
Duration (sec)	1.30	1.14	<1

* $p < .05$, one-tailed.

General Discussion

Four-month-old infants can perceive a bimodally specified episode when they detect a temporal invariance in light and sound. Under some conditions, infants explore this episode by looking and listening. Like human adults, human infants appreciate that optic and acoustic stimulation sometimes provide information for one event and sometimes do not.

Infants are sensitive to temporal invariants of at least two kinds. First, they can detect the common rate of sound bursts and the visible impacts of an object and surface, even if the sounds and impacts are not simultaneous. Second, infants can detect the simultaneous occurrence of sound bursts and visible impacts, even if the tempo of the auditory accompaniment accords with the movements of both the appropriate and the inappropriate objects. There are limits, no doubt, to the ability to detect such relationships. A temporal invariance of sound and motion may escape an infant's notice if an event is sufficiently complex. Furthermore, perceivers of any age will surely fail to detect the simultaneity of sounds and impacts if an object oscillates too rapidly, and they will not detect the common rate of sounds and impacts if it moves too slowly. Despite these limitations, infants should be able to detect temporally invariant information for a variety of natural events. Babies may perceive an auditory–visual relationship when they view a percussion sequence, a game of pat-a-cake, or the movements of a toy (cf. Spelke, 1976; Bahrick et al., Note 3) on this basis.

Although these experiments provide evidence that infants can perceive a bimodally specified event through a process of invariant detection, they do not rule out the possibility that bimodal perception can be achieved in other ways as well. Infants may sometimes come to perceive the unity of an audible and visible episode through processes of reciprocal assimilation or association; these processes may operate in situations that have been untested so far. Furthermore, an ability to learn by association could have contributed to infants' performance in two of the present studies. The results of Experiments 1 and 3 are consistent with one specific version of an associationist theory. When infants confront a bouncing, sounding object, they might segment the sound stream into bursts, segment the visible event into periods of impact and nonimpact (times when the object sits on the ground or dangles in the air), and associate the onset of each sound burst with each moment of visible impact. Further research might test specific association and assimilation hypotheses directly. If such studies are conducted with sufficiently young infants, they should ultimately reveal how infants first discover the unity of a bimodally specified event.

In summary, these experiments support an invariant-detection description of infant perception. Young babies do not appear to experience a world of unrelated visual and auditory sensations. They can perceive unitary audible and visible events. Infants can perceive the unity of a moving, sounding object that they see for the first time by detecting a temporal relation between the object's sound and its visible movement. This ability helps infants to explore events by looking and listening. Thus, it may lead them to discover other stimulus relationships that unite what they see with what they hear.

Reference Notes

1. Gibson, E. J. *The ecological optics of infancy: The differentiation of invariants given by optical motion.* Presidential address at the meeting of the American Psychological Association, San Francisco, August 1977.
2. Lewis, M., & Hurowitz, L. *Intermodal person schema in infancy: Perception within a common auditory-visual space.* Paper presented at the meeting of the Eastern Psychological Association, Boston, April 1977.
3. Bahrick, L., Walker, A., & Neisser, U. *Infants' perception of multimodal information in novel events.* Paper presented at the meeting of the Eastern Psychological Association, Washington, D.C., March 1978.
4. Spelke, E. *The infant's acquisition of knowledge of bimodally specified events.* Unpublished manuscript, 1978. (Available from E. Spelke, Department of Psychology, University of Pennsylvania, Philadelphia, Pa. 19104.)

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.

Berkeley, G. Essay towards a new theory of vision, 1709. In A. C. Fraser (Ed.), *Selections from Berkeley*. Oxford, England: Clarendon Press, 1910.

Birch, H., & Lefford, A. Visual differentiation, intersensory integration, and voluntary motor control. *Monographs of the Society for Research in Child Development*, 1967, *32*(2, Serial No. 110).

Blank, M., & Bridger, W. H. Cross-modal transfer in nursery school children. *Journal of Comparative and Physiological Psychology*, 1964, *58*, 272–282.

Butterworth, G., & Castillo, M. Coordination of auditory and visual space in newborn human infants. *Perception*, 1976, *5*, 155–160.

Fieandt, K., & Gibson, J. J. The sensitivity of the eye to two kinds of continuous transformation of a shadow pattern. *Journal of Experimental Psychology*, 1959, *57*, 344–347.

Field, J., DiFranco, D., Dodwell, P., & Muir, D. Auditory-visual coordination in 2½-month-old infants. *Infant Behavior and Development*, 1979, *2*, 113–122.

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

Gibson, E. J., Owsley, C. J., & Johnston, J. Perception of invariants by 5-month-old infants: Differentiation of two types of motion. *Developmental Psychology*, 1978, *14*, 407–415.

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

Gibson, J. J. *An ecological approach to visual perception.* Boston: Houghton-Mifflin, 1979.

Haith, M. M. Visual scanning in infants. In L. J. Stone, H. T. Smith, & J. B. Murphy (Eds.), *The competent infant.* New York: Basic Books, 1973.

Horowitz, F. D. Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development*, 1974, *39*(5–6, Serial No. 158).

Lyons-Ruth, K. Bimodal perception in infancy: Response to auditory-visual incongruity. *Child Development*, 1977, *48*, 820–827.

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*(4, Serial No. 167).

Mill, J. *Analysis of the phenomena of the human mind.* London: Longmans & Dyer, 1829.

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

Piaget, J. *The origins of intelligence in children.* New York: International Universities Press, 1952.

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., & Owsley, C. J. Intermodal exploration and knowledge in infancy. *Infant Behavior and Development*, 1979, *2*, 13–27.

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

Received March 30, 1979 ■