Developmental Psychology
1987, Vol. 23, No. 5, 665–670

0012-1649/87/$00.75

Effect of Auditory Numerical Information on Infants' Looking Behavior: Contradictory Evidence

David Moore, Joyce Benenson, J. Steven Reznick, Michele Peterson, and Jerome Kagan
Harvard University

Seven-month-old infants saw pairs of slides of two and three objects while listening to either two or three drum beats. Subjects looked longer at the noncorresponding display, particularly during the second block of 16 trials. These results are not in accord with those of Starkey, Spelke, and Gelman (1983). Although 7-month-old infants' attentiveness to a visually presented numerical event might be influenced by auditory information, the data call into question the suggestion that this influence is mediated by crossmodal matching of numerical information.

Research on intermodal perception in infancy has led some investigators to claim that very young infants can extract a schema in one modality and recognize that schema in another modality (Aronson & Rosenbloom, 1971; Bahrick, 1983; Meltzoff & Borton, 1979; Spelke, 1976; Wagner, Winner, Cicchetti, & Gardner, 1981), even though this research has not gone unchallenged. McGurk and Lewis (1974) present data that "afford no support for the hypothesis that the very young human infant lives in a perceptually unified audiovisual world" (p. 650), and, although Bushnell's work (Bushnell, 1982; Lockman, Ashmead, & Bushnell, 1984) suggests that infants do develop intermodal perception, it does not seem to appear until they are at least 9 months of age. Although evidence that young children possess a capacity to understand the concept of number has been accumulating for more than 15 years (Antell & Keating, 1983; Gelman & Gallistel, 1978; Klahr & Wallace, 1976; Mehler & Bever, 1967; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1983; Strauss & Curtis, 1981), the claim that 7-month-old infants can extract an abstract dimension such as number across two different sensory modes is original and theoretically important (Starkey et al., 1983). The study reported in this article summarizes an attempt to replicate this result.

The study was designed to assess the ability of 6- to 8-month-old infants to detect numerical correspondences between visible and audible arrays, using differential attentiveness as the dependent variable, in a procedure similar to that used by Spelke (1976) and, except for the differences to be noted later, identical with that used by Starkey et al. (1983).

Method

Subjects

The subjects were 26 infants (15 boys and 11 girls) between 6 months, 6 days and 7 months, 27 days of age (mean age = 6 months, 28 days). All infants were healthy, and 24 of the 26 were full-term. Each subject was volunteered by a caretaker who brought the infant to the laboratory and was present with the infant throughout the procedure.

Stimuli

The 16 pairs of visual stimuli were produced by photographing small household objects against a white background. Each pair included one slide of 2 objects and one slide of 3 objects. The first 8 pairs of slides contained a total of 40 ($[2 \times 8] + [3 \times 8]$) different objects. The second 8 pairs of slides used 32 of the same objects, plus 8 new objects. The 3-object slides of this second set were created by photographing each pair of objects that was originally used as a 2-object configuration with 1 of the 8 new objects. The 2-object slides of the second set were created by removing one of the original objects from each of the 3-object configurations of the first 8 pairs of slides. No pair of slides presented simultaneously in a trial contained any identical objects. In addition, the spacing and positioning of the objects over the background was deliberately varied across slides to render these variables irrelevant to looking behavior. An attempt was also made to control for the amount of interest the slides would elicit, by subjectively matching each pair for perceptual salience. Objects in both slides of each pair were chosen to be of similar colors, sizes, and complexities. The composition of the slides is presented in Table 1 in an actual order seen by one of the subjects.

The auditory stimuli, recorded on one track of a ¼ in. reel-to-reel tape, consisted of sets of two and three drum beats. Each beat occurred approximately 1.33 s after the last beat, regardless of whether a two- or a three-beat set was presented; hence, the presentation of a three-beat set required more time than the presentation of a two-beat set (Starkey et al. [1983] explored the effects of a duration discrepancy such as this one and reported that it did not account for their findings). The stimulus tape was created in the following way. One set of two beats and one set

Preparation of this article was made possible through support from John D. and Catherine T. MacArthur Foundation and National Institute of Mental Health Grant MH 40619.

We wish to thank anonymous reviewers for suggesting our title and for their very helpful comments on the text. An earlier version of this article was presented at the Fifth International Conference on Infant Studies, Los Angeles, California, April 1986.

Correspondence concerning this article should be addressed to David Moore, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138.

Table 1
Order of Displays of Visual Stimuli Given to One Infant

Trial	Objects seen: Left	Objects seen: Right	No. of drumbeats
1	Orange Helicopter	Hole-puncher Sewing kit Yellow spoon	2
2	Blue lego Red pot	Lion Recorder cleaner Bell	2
3	Red truck Blue cup	Chalkboard eraser Scotch tape Toothbrush	2
4	Toy lizard Red robot	Potato Scissors Green block	2
5	Measuring spoon Apple	Red ashtray Banana Blue ribbon	3
6	Toothbrush case Blue ball	Gray bird Pear Fork	3
7	Purple triangle Yellow stick	Green car Red rectangular block Blue cylinder	3
8	Purple spool Keys	Christmas tree Red hourglass block Blue block	3
9	Measuring spoon Apple Yellow corkscrew	Red ashtray Banana	2
10	Blue lego Red pot Yellow stick	Lion Recorder cleaner	2
11	Yellow triangle Red truck Blue cup	Chalkboard eraser Scotch tape	2
12	Orange Helicopter Green crayon	Hole-puncher Sewing kit	2
13	Toy lizard Red robot Crayon box	Potato Scissors	3
14	Red pitcher Toothbrush case Blue ball	Fork Pear	3
15	Purple triangle Yellow stick Gold car	Green car Red rectangular block	3
16	Yellow car Purple spool Keys	Christmas tree Red hourglass block	3

of three beats were recorded at constant volume, with the tempo of the beats guided by a stopwatch. These recordings were copied repeatedly to form the stimulus tape and to ensure that the drum beats were identical on all trials. The tape was constructed so that four trials with two drum beats were followed by four trials with three drum beats, and this pattern was repeated throughout the experiment. Trials were separated by a 4-s intertrial interval. The second track of the audio-tape contained signals for a computer that operated the slide projectors.

Procedure

Each caretaker first signed an informed consent form and filled out a questionnaire requesting general information about the child. The subject was then seated in the caretaker's lap, 3 ft from a rear projection screen in a dimly lit testing room. Immediately above the screen was a small window through which the entire session could be observed and videotaped.

Each caretaker was given some general information concerning the procedure and was asked to orient the infant toward the screen and keep the infant's body centered with respect to the screen. Caretakers were also told to do whatever was necessary to keep the infant in a quiet, alert state, as long as this action would not interfere with the infant's viewing of the slides. Finally, caretakers were told to keep their eyes closed so that they could not systematically influence the looking behavior of the infants. This request posed no problem for the caretakers, who were given a chance to see all of the slides at the conclusion of the test session.

A trial began with the projection of one pair of slides on the screen in front of the subject. The slides subtended 44.3° × 31.5° of visual angle, separated by 5.7° (3 ½ in.) of blank screen. Approximately 1.25 s later, the subject heard a set of either two or three drum beats coming from an audio speaker located centrally behind the projection screen. The slides remained on the screen during the auditory presentation and for 8 s after the termination of the drum beats. A single two- or three-beat sequence was presented on each trial. There were two blocks of 16 trials each, the second block being an exact repetition of the first block. Subjects who were irritable or fatigued during or at the end of the first block were given a break before the second block of trials began. The break lasted up to 8 min, during which the caretaker took the infant for a short walk.

Twelve subjects heard two drum beats accompany a particular pair of slides; the remaining 14 infants heard three drum beats accompany the same pair of slides. The lateral position of any particular slide was counterbalanced across subjects, so that 14 subjects saw a particular slide on the left, and the other 12 saw the same slide on the right. Each subject experienced eight trials per block in which the two-object slide was presented on the right, and eight trials per block in which the two-object slide was presented on the left. Finally, two different orders of slide presentation were used, with one-half of the subjects viewing each order.

Duration of looking at each of the two slides was recorded during each trial by one of two experimenters trained to record infants' eye movements. Intercoder reliability for fixation time, assessed with a subset of subjects, was .90. The duration of each fixation was stored in a computer, along with information concerning the time when a subject looked at a particular slide (i.e., during the 1.25 s before the onset of the drum beats, during the presentation of the auditory stimuli, or during the 8-s test period following this presentation). Finally, each subject's emotional state for each trial was recorded by a different observer so that trials in which a subject was fretful could be eliminated from the statistical analyses. The study was essentially a replication of Starkey et al. (1983), except for arguably inconsequential methodological differences in the durations of trials, intertrial intervals, and interblock intervals, and in the pitch, brightness, and variability of the stimuli. In addi-

Table 2
Attention to Numerically Corresponding Versus Noncorresponding Displays, by Number of Drum Beats and by Block, Before Data Deletions

Trial block	No. of drum beats	Mean fixation corresponding display (in s) *M*	*SD*	Mean fixation noncorresponding display (in s) *M*	*SD*	Mean fixation of corresponding display[a] (%)	*t*	Mean proportion of subjects looking at corresponding display[b] (%)	*n*
1 & 2	2 & 3	1.78	0.65	2.16	1.02	47**	2.24	37	24
1	2 & 3	2.02	0.85	2.12	1.11	49	0.55	39	23
1	2	1.46	0.77	2.31	1.58	44	1.62	39	23
1	3	2.55	1.27	1.97	0.86	53	1.40	61	23
2	2 & 3	1.43	0.88	2.03	1.14	43**	2.57	27*	22
2	2	1.31	0.87	2.14	1.09	44	1.72	30	20
2	3	1.65	1.30	2.01	1.46	43*	2.08	19***	21

[a] $P_d = D_c/(D_c + D_n)$, where P_d is the mean proportion of duration of fixation averaged over the sets of corresponding displays (c) and noncorresponding displays (n). This proportion was compared with that expected by chance (0.50); significance was assessed by *t* tests.
[b] $P_s = S_c/(S_c + S_n)$, where P_s is the proportion of subjects and S_c and S_n are the numbers of subjects whose mean proportion of duration was greater on the corresponding displays (c) or the noncorresponding displays (n); significance was assessed by two-tailed sign tests.
* $p < .052$. ** $p < .05$. *** $p < .01$.

tion, although neither the coders in this study nor Starkey's coders were "deaf" to the auditory stimuli, Starkey's coders were blind. Because of the configuration of our rear projection screen, observers in our study could have stepped back and looked at the visual stimuli, although they reported that they did not do this.

Results and Discussion

A total of 730 trials were collected across all subjects. Eighty-two trials were excluded from the analyses because of infant fretfulness (crying through a number of trials) or technical problems. Therefore, all initial analyses were conducted on a maximum of 648 trials. The various initial analyses summarized below were based on a maximum of 24 and a minimum of 20 subjects, who contributed a maximum of 32 trials and a minimum of 1 trial to the analyses.

The preliminary analyses, parallel to those conducted by Starkey et al. (1983), generated statistics comparable across both studies. In addition, an analysis of variance (ANOVA) was conducted to test for all main effects and interactions associated with sex, trial blocks, and number of beats. All analyses used the subjects' total fixation time after termination of the auditory stimulus as the dependent variable. Following Starkey et al., the average percentage of time subjects looked at the corresponding display (i.e., the three-object visual display when the subjects heard three drum beats and the two-object visual display when the subjects heard two drum beats) was calculated across trials and compared with the percentage expected by chance (.50) using two-tailed *t*-tests. In addition, the proportion of subjects who looked more at the corresponding slide, collapsed across their trials, was assessed with a two-tailed sign test. The results of these analyses appear in Table 2.

When the data were collapsed across both number of beats and blocks, subjects looked significantly longer at the display that did not correspond to the number of beats in the auditory stimulus, $t(23) = 2.24$, $p < .05$. This effect is the result of the subjects' preference for the noncorresponding display in the second block of trials.

We reanalyzed the data, deleting trials (and/or subjects) that may have been biased for a number of reasons. First, data from an additional 172 trials were eliminated because infants did not look at both slides. Second, because it may be misleading to treat the scores of subjects who provided only a few trials as equivalent to the scores of those who contributed data for a majority of trials, all subjects who completed less than 40% of the total possible number of trials in a particular analysis (after the deletions noted above) were removed from that analysis. Data were also eliminated similarly if an infant showed a visual preference for one side (defined as looking to that side over 80% of the subject's total looking time summed across all trials). Table 3 reveals the number of subjects excluded from each analysis under each of these criteria. The final analyses to be summarized were based on a maximum of 24 and a minimum of 12 subjects who contributed a maximum of 30 and a minimum of 5 trials to the analyses.

These deletions did not substantially change the results of the original analysis. Although there was no evidence that looking patterns were significantly influenced by the auditory stimuli when the data were collapsed across blocks, during the second block more subjects (12 out of 13, $p < .01$) looked at the noncorresponding display significantly longer than at the corresponding display, $t(12) = -4.6$, $p < .001$ (see Table 4).

The data were also analyzed in a 2 × 2 × 2 (Beats × Blocks × Sex) repeated measures ANOVA, using percentage of time infants looked at the corresponding display as the dependent variable. This analysis revealed no additional findings; the only significant result was a main effect of block, so that subjects' preferences (i.e., fixation times) for noncorresponding displays were stronger in Block 2 than in Block 1, $F(1, 19) = 6.86$,

Table 3
Number of Subjects Deleted by Analysis and by Criterion

Level of analysis	Reason for data deletion[a]: Insufficient data after initial deletions	Visual side preference (total)	Visual side preference (new)[b]
Across blocks & across beats	5	0	0
Across blocks & by beats	12	5	3
By blocks & across beats	10	3	1
By blocks & by beats	19	9	5

[a] See Results and Discussion section for definitions of *insufficient data* and *visual side preference.*
[b] Number of subjects deleted due to a visual side preference who were *not* deleted from analyses due to insufficient data.

$p < .01$. A contrast analysis designed to test the hypothesis suggested by the t-tests just described was significant; subjects preferred the noncorresponding display in Block 2 during both two- and three-beat trials, and in Block 1 when they were hearing two drum beats, $F(1, 19) = 6.20$, $p < .05$.

An additional Beats × Block × Sex repeated measures ANOVA was implemented with percent of time subjects preferred the three-object display in each trial as the dependent variable. In contrast to the dependent variable reported above, percent of time preferring three objects is based on the category of stimulus studied, independent of the number of drum beats heard. Although this analysis revealed no additional findings, it recast the data into a more easily interpreted form.

Consistent with the results reported above, there was a main effect of beat, $F(1, 22) = 5.19$, $p < .05$, and a Beats × Block interaction, $F(1, 16) = 6.16$, $p < .05$. As shown in Figure 1 and confirmed by a post hoc Duncan New Multiple Range test, infants looked longer at three objects than at two objects, except when they heard three drum beats in the second block.

The tendency to look longer at three- than at two-object displays is in accord with the findings of Karmel (1969), McCall and Kagan (1967), and Maisel and Karmel (1978). Fantz and Fagan's (1975) suggestion that there is a positive relation between contour density in a visual display and an infant's fixation time is a plausible explanation of this result (Karmel, 1969; McCall & Kagan, 1967).

But why were two objects preferred by infants hearing three beats during Block 2? Because a block effect was present in both these results and those reported by Starkey et al. (1983), the data were analyzed again after being divided into four 8-trial segments. It was hoped that this analysis would offer some insight into this unexpected result. This analysis revealed a main effect of beats, $F(1, 22) = 5.37$, $p < .05$, so that subjects preferred to look at three objects more when they heard two beats. There was also a main effect of trial segment, $F(1, 22) = 9.85$, $p < .01$, so that subjects' preferences for three-object displays increased over trials within each block, regardless of the number of drum beats heard. In addition, this analysis revealed a Beat × Block interaction, $F(1, 16) = 8.64$, $p < .01$, and a Beat × Block × Trial Segment interaction, $F(1, 13) = 7.48$, $p < .01$. The results of this analysis are illustrated in Figure 1.

During Block 1, subjects looked longer at displays containing three objects, regardless of the number of beats they were hearing, and this preference increased over the 16 trials of Block 1. After returning to the testing situation following a short break, subjects looked longer at displays containing two

Table 4
Attention to Numerically Corresponding Versus Noncorresponding Displays, by Number of Drum Beats and by Block, After Data Deletions

Trial block	No. of drum beats	Mean fixation corresponding display (in s) M	SD	Mean fixation noncorresponding display (in s) M	SD	Mean fixation of corresponding display[a] (%)	t	Mean proportion of subjects looking at corresponding display[b] (%)	n
1 & 2	2 & 3	2.10	0.61	2.37	0.83	48	1.39	37	19
1	2 & 3	2.18	0.78	2.39	0.96	48	0.99	38	21
1	2	1.62	0.61	2.79	1.46	40*	2.47	22*	18
1	3	2.63	1.05	2.28	0.83	53	1.20	57	21
2	2 & 3	2.18	0.61	2.97	0.70	42***	4.61	8**	13
2	2	1.87	0.83	3.12	1.31	41	1.82	25	12
2	3	2.04	1.11	2.75	1.18	39*	2.43	33	12

[a] $P_d = D_c/(D_c + D_n)$, where P_d is the mean proportion of duration of fixation averaged over the sets of corresponding displays (c) and noncorresponding displays (n). This proportion was compared with that expected by chance (0.50); significance was assessed by t tests.
[b] $P_s = S_c/(S_c + S_n)$, where P_s is the proportion of subjects and S_c and S_n are the numbers of subjects whose mean proportion of duration was greater on the corresponding displays (c) or the noncorresponding displays (n); significance was assessed by two-tailed sign tests.
* $p < .05$. ** $p < .01$. *** $p < .001$.

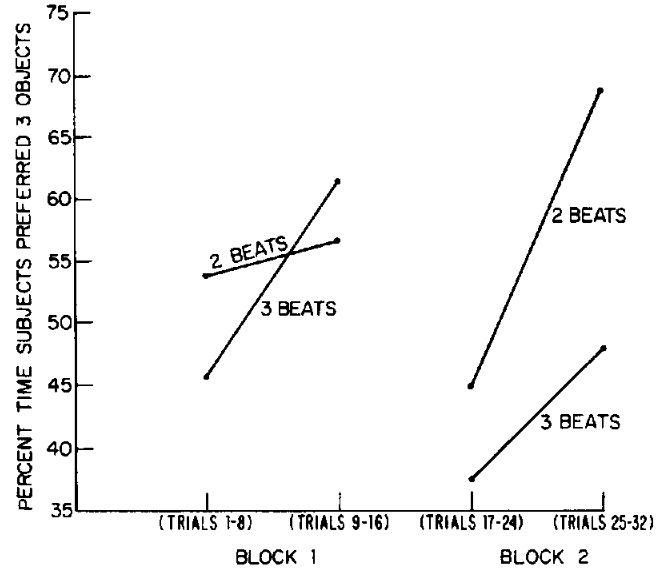

Figure 1. Percentage fixation time to Three-Object Displays over trial blocks.

objects. But, as the second block of 16 trials proceeded, the subjects once again looked longer at displays of three objects. In addition, by the end of this block, there was an effect of the drum beats; subjects hearing two drum beats looked longer at the displays of three objects than did subjects hearing three drum beats.

This description was confirmed using post hoc Duncan New Multiple Range tests at a sensitivity level of .05. Specifically, there was a significant percentage increase in fixation of three objects across the 16 trials of each block for three-beat trials in both blocks and for two-beat trials in the second block. There was a significant decrease in percent fixation of three objects from Trial Segment 2 in Block 1 to Trial Segment 1 in Block 2 for both two- and three-beat trials. Finally, the percentage of preference for three objects was significantly greater when subjects heard two beats than when they heard three beats, but only in the second trial segment of Block 2.

It is not obvious why subjects' fixation times to displays of three objects become longer over trials. One possible explanation holds that, initially, subjects may have been too preoccupied with the novel auditory stimulus to attend carefully to the visual stimuli. This dilution of attention could have produced the apparently random looking behaviors seen during the initial few trials. If the infants gradually devoted increasing attention to the visual stimuli, their natural tendency to prefer displays containing more contour would have led them to look longer at three-object displays during the remaining trials of the first block.

When the subjects returned from the short break between blocks, however, they preferred displays of two objects to those with three objects. This unexpected preference for two objects may reflect habituation to three objects and dishabituation to the novelty of two objects. That is, subjects looked more at three-object displays than at two-object displays during the previous block, and were probably bored; thus, they looked more at displays of two objects in the next set of eight trials.

It is not clear, however, why the auditory stimuli became influential during the last eight trials of the second block. The late appearance of a drum-beat effect may be the result of the subjects' inability to discriminate two from three drum beats until they had gained some experience with these stimuli. Although it is not known how such a discrimination could be learned, such gradual learning would account for the effect reported here as well as the block effect reported by Starkey et al. (1983).

The data suggesting a preference for noncorresponding over corresponding displays in the last half of the second block of trials could be accounted for in a number of ways. Infants might be demonstrating a novelty preference in which they look at a visual display containing a novel number of elements, compared with the auditory event. This interpretation would support Starkey's contention that infants are able to detect numerical information, but it is unclear how subtle methodological differences could cause infants to prefer a novel numerical stimulus in this experiment and a familiar numerical stimulus in Starkey's research. An alternative interpretation is suggested by Turkewitz (Lawson & Turkewitz, 1980; Lewkowicz & Turkewitz, 1981): Infants are biologically prepared to seek an optimal level of stimulation across all sensory modes. Infants who hear three beats may be "driven" to visually explore the less stimulating display of two objects, whereas infants who hear two beats are driven to explore the more stimulating display of three objects. According to this interpretation, the differences between our results and those of Starkey et al. (1983) may be due to undocumented differences in the loudness and/or the pitch of the auditory stimuli and the size and/or brightness of the visual stimuli.

It must be noted that these interpretations are for an effect that was present only during the final 8 trials of the 32-trial experiment, and that the data were exactly opposite to the effect reported by Starkey et al. (1983). We recognize that this effect is fragile and might be attributable to chance. Our data *do* reveal infants' preferences for greater contour density and their tendency to habituate to repeatedly presented stimuli. They *might* reveal that 6- to 8-month-old infants' attentiveness to visual events is influenced by information received aurally. However, they *do not* support the claim that infants prefer to look at arrays that correspond to a specific number of sounds and, therefore, question the robustness of the Starkey et al. result.

References

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development. 54,* 695–701.

Aronson, E., & Rosenbloom, S. (1971). Space perception in early infancy: Perception within a common auditory–visual space. *Science, 172,* 1161–1163.

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

Bushnell, E. W. (1982). Visual–tactual knowledge in 8-, 9½-, and 11-month-old infants. *Infant Behavior and Development, 5,* 63–75.

Fantz, R. L., & Fagan, J. F., III. (1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child Development, 16,* 3–18.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA: Harvard University Press.

Karmel, B. Z. (1969). The effect of age, complexity, and amount of contour density on pattern preferences in human infants. *Journal of Experimental Child Psychology, 7,* 339–354.

Klahr, D., & Wallace, J. G. (1976). *Cognitive development: An information processing view.* Hillsdale, NJ: Erlbaum.

Lawson, K. R., & Turkewitz, G. (1980). Intersensory function in newborns: Effect of sound on visual preferences. *Child Development, 51,* 1295–1298.

Lewkowicz, D. J., & Turkewitz, G. (1981). Intersensory interaction in newborns: Modification of visual preferences following exposure to sound. *Child Development, 52,* 827–832.

Lockman, J. J., Ashmead, D. H., & Bushnell, E. W. (1984). The development of anticipatory hand orientation during infancy. *Journal of Experimental Child Psychology, 37,* 176–186.

Maisel, E. B., & Karmel, B. Z. (1978). Contour density and pattern configuration in visual preferences in infants. *Infant Behavior and Development, 1,* 127–140.

McCall, R. B., & Kagan, J. (1967). Attention in the infant: Effects of complexity, contour, perimeter, and familiarity. *Child Development, 38,* 939–952.

McGurk, H., & Lewis, M. (1974). Space perception in early infancy: Perception within a common auditory-visual space? *Science, 186,* 649–650.

Mehler, J., & Bever, T. G. (1967). Cognitive capacity of very young children. *Science, 158,* 141–142.

Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature, 282,* 403–404.

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

Starkey, P., & Cooper, R. G., Jr. (1980). Perception of numbers by human infants. *Science, 210,* 1033–1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179–181.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development, 52,* 1146–1152.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development, 52,* 728–731.

Received October 7, 1985
Revision received December 22, 1986
Accepted January 5, 1987 ■