Developmental Psychology
1997, Vol. 33, No. 3, 423–428

0012-1649/97/$3.00

Numerical Abstraction in Infants: Another Look

Kelly S. Mix, Susan Cohen Levine, and Janellen Huttenlocher
University of Chicago

This article examines an important finding from the literature on infant numerical competence. The finding, reported by P. Starkey, E. S. Spelke, and R. Gelman (1990), was that infants looked longer toward a visual display that was equal in number to an auditory set. In Experiment 1, when the procedures described by P. Starkey et al. were followed and duration was held constant across auditory sequences that varied in number, infants looked longer toward the display that was not numerically equivalent to the auditory set. In Experiment 2, when the rate and duration of the auditory sequences were varied randomly within infants, no significant preference for either the equivalent or nonequivalent visual display was shown. These results raise questions about P. Starkey et al.'s claims that infants can represent the numerosity of sets in different modalities and then perform one–one correspondence computations over them.

Interest in the origins of number concepts has led to several studies focusing on numerical competence in infants. Most of these have evaluated the ability of infants to discriminate between small set sizes by using the habituation procedure. In these studies, it was found that infants' looking time decreased after being shown several arrays with the same small number of items (usually two or three) but that looking time recovered when a novel number of items was shown (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981). The same result was found when experimenters in these studies varied the displays to control for differences in brightness, density, line length, contour, area, and homogeneity of set items, indicating that the infants used numerical information, rather than these other factors, as the basis for making discriminations.

Although the habituation studies provided evidence of a sensitivity to numerical information during infancy, they gave little indication of what particular process might underlie the infants' performance. Starkey, Spelke, and Gelman (1990) argued that although infants might possess central numerical processes, such as one–one correspondence, this had not been clearly demonstrated because the use of other processes (e.g., use of configural information or visual subitizing) could explain infants' looking preferences in the visual habituation experiments. They proposed that if infants could recognize numerical equivalence between visual displays and sounds, this would demonstrate an ability to establish one–one correspondence because these other processes could not be applied to temporally distributed sets.

To test whether infants could recognize auditory–visual numerical correspondences, Starkey et al. (1990) designed a series of experiments in which 6- to 8-month-olds were shown pairs of visual displays that included one display of two objects and one display of three objects. While the displays were still visible, infants heard either two or three drumbeats. Measurement of looking time revealed that the infants looked significantly longer toward the display that matched the number of sounds. Starkey et al. interpreted these results as evidence that infants can perceive the number of distinct entities both in a sequence of sounds and a visual display and can relate these sets to one another in terms of numerical equivalence. They contended that to detect such relations, infants must make use of a process involving both one–one correspondence and the abstraction principle (i.e., knowledge that any discrete element, including sounds, can be enumerated; Gelman & Gallistel, 1978). They further proposed that the emergence of these abilities is dependent on neither the acquisition of language nor a culture-specific counting system.

This interpretation attributes to infants a competence far beyond that shown in previous studies of infant number concepts. As such, it has had a strong influence on theories of numerical development (e.g., Gelman & Brenneman, 1994; Starkey, 1992; Wynn, 1992). However, there are several reasons that the claims advanced by Starkey et al. (1990) warrant further investigation. First, the results they reported for infants have not been replicated. In the only replication attempt to date (Moore, Benenson, Reznick, Peterson, & Kagan, 1987),[1] infants showed a significant preference for one of the visual displays but in the opposite direction of that reported by Starkey et al. That is, infants looked longer at the display that was not equivalent to the number of sounds. Starkey et al. attributed the discrepancy between their findings and those of Moore et al. to differences in statistical

Kelly S. Mix, Department of Psychology, University of Chicago; Susan Cohen Levine and Janellen Huttenlocher, Departments of Psychology and Education, University of Chicago.

The study presented in this article was supported by a grant from the University of Chicago School Mathematics Project. Portions of this research were presented at the annual meeting of the Midwestern Psychological Association, Chicago, Illinois, May 1994. We gratefully acknowledge Rebecca Compton, Katchen Locke, and Kirste Moline for their assistance with data collection and data entry.

Correspondence concerning this article should be addressed to Kelly S. Mix, who is now at the Department of Psychology, Indiana University, Bloomington, Indiana 47405. Electronic mail may be sent via Internet to kmix@indiana.edu.

[1] This replication attempt was conducted in response to an earlier publication that reported a portion of the studies described in Starkey et al. (1990; see also Starkey, Spelke, & Gelman, 1983).

analysis, such as Moore et al.'s removing from the data set infants who failed to inspect both displays, and differences in experimental procedure, such as Moore et al.'s allowing irritable or fatigued infants to take breaks of up to 8 minutes between trial blocks. However, if these differences could lead to a complete reversal in preference, it suggests that the effect of looking toward the numerically equivalent display in an auditory–visual matching task is fragile and may be sensitive to a variety of factors.

A second reason to investigate the claims made by Starkey et al. (1990) is that the level of sophistication attributed to infants in their interpretation surpasses that which has been shown in preschool children. Mix, Huttenlocher, and Levine (1996) investigated whether 3- and 4-year-olds showed evidence of the ability to detect auditory–visual numerical correspondences. The study used two matching task conditions. In the auditory–visual condition, children heard a set of sounds and then indicated which of two visual arrays was numerically equivalent to the sounds. In the visual–visual condition, children were first shown a set of objects. Then the set was covered and children chose between the visual arrays. (The same visual array pairs were used in both the auditory–visual and visual–visual conditions.) It was found that 3-year-olds performed at chance in the auditory–visual condition, even though they successfully completed the visual–visual matching task. In contrast, 4-year-olds performed significantly above chance in both conditions, indicating that the ability to detect auditory–visual numerical correspondences develops during this age period. Furthermore, Mix et al. found evidence that success on the auditory–visual matching task was related to mastery of the conventional counting system. These results are surprising given the claims of infant competence made by Starkey et al.

Because the experimental findings and interpretations advanced by Starkey et al. (1990) are of far-reaching importance for theories of numerical development, further investigation is needed. In the present study, we first attempt to test the robustness of the Starkey et al. findings, and then we introduce an important variation in procedure.

Experiment 1

This experiment used the same procedures described by Starkey et al. (1990), with two exceptions. First, Starkey et al. used aerial photographs of household objects for visual displays. We chose instead to use black dots on a white background because this eliminated potentially distracting features of the visual entities that might interfere with infants' ability to recognize numerical relations. In this way, we aimed to simplify the infants' task by making the numerical information in the sets more salient. The second difference was in the order of presentation of the drumbeat sequences. In Starkey et al.'s procedure, the numerosity of the sequences was blocked, so for example, 8 trials with two drumbeats were presented followed by 8 trials with three drumbeats. In the present experiment, the number of drumbeats presented in the auditory sequences was randomized across the 32 trials rather than blocked. This tested whether infants could detect numerical equivalence on a trial-by-trial basis as the Starkey et al. interpretation would predict.

Method

Participants. Twenty healthy full-term infants (8 boys and 12 girls) participated in the experiment. They ranged in age from 6 months 3 days to 8 months 23 days (M = 7 months 6 days). Five additional infants were excluded because they either cried or fell asleep during the experiment.

Stimuli. The visual displays were slides of two or three black dots on a white background. The dots were positioned in random configurations on an imaginary 3 × 3 matrix.[2] The separation between individual dots in a given display was between 2° and 8° of visual angle. The slides were rear-projected onto two screens that subtended 38° × 11° of visual angle, separated by 17° of blank space. The front of the wooden apparatus containing the screens was painted black and measured 5 × 5 ft (1.5 × 1.5 m). Slides were presented in pairs that included one slide of two dots and one slide of three dots. For each infant, the side on which a particular numerosity appeared was changed after the first half of trials was presented.

The auditory stimuli were produced by beating a biscuit tin with a wooden stick. The tin was located approximately 5 ft (1.5 m) behind the projection screen and was centered between the two visual displays. The experimenter who produced the drumbeats was guided by a counter that appeared on a computer monitor. The counter displayed numbers at a rate of one per second and helped ensure that the presentation of the drumbeats was consistent throughout the experiment. Because the overall duration of the drumbeat sequences was held constant, the tempo of the two-drumbeat sequence was slower than the tempo of the three-drumbeat sequence. The three-drumbeat sequence was presented at a rate of one beat per second, and the two-drumbeat sequence was presented at a rate of one beat per 1.5 s (i.e., one beat on both the first and third seconds of a 3-s interval).

Procedure. Infants were seated in a caretaker's lap, 3 ft (0.9 m) from the projection screen. Caretakers wore a blindfold that prevented them from viewing the visual displays and possibly biasing the infants' looking behavior. Each trial began with the presentation of one pair of slides, followed approximately 1 s later by the onset of the drumbeat sequence. The display pair remained visible for 10 s following the termination of the drumbeat sequence. The slides were then forwarded to reveal the display pair for the next trial.

There were two blocks of 16 trials each. Each infant was presented with at least one complete block. The entire second trial block also was presented unless an infant became fussy. Sixteen infants completed both trial blocks, receiving a total of 32 trials each. The remaining 4 infants completed 27–31 trials. One of these 4 infants completed both blocks, but his first trial was lost because of experimenter error. Infants were not given breaks between trial blocks.

There were 32 auditory sequences, half with two drumbeats and half with three drumbeats. These were presented in a fixed random order. The slides were presented in one of two orders that counterbalanced the lateral positions of each numerosity and the number of drumbeats accompanying a given display pair across infants. Gender was controlled by assigning half of the boys and half of the girls to each presentation order.

The duration of looking time toward each member of a display pair was recorded by one of two observers who viewed the infant through a peephole located between the two displays. Observers were trained to detect eye movements toward any part of the display through practice with both adults and infants. They pressed one of three keys on a computer keyboard to indicate looking toward the left or right side of the display, as well as looking toward neither side. The duration of each

[2] Through random chance, the majority of the displays with three dots formed triangles (30 out of 32); the remaining two displays of three dots formed lines, and, of course, all displays of two dots formed lines.

Table 1
Looking at Equivalent and Nonequivalent Displays in Experiment 1

	Mean duration of looking (s)		Proportionate effect of sound[a]		
Trial block	Equivalent display	Nonequivalent display	Two-dot display	Three-dot display	Overall
Block 1	2.59	2.73	.48	.50	.49
Block 2	1.95	2.21	.46	.46*	.46
Blocks 1 and 2	2.27	2.47	.46*	.48*	.47**

[a] Proportionate effect of sound refers to the percentage of looking time that infants attended to a given display when the equivalent number of drumbeats was played.
$*p < .05$. $**p < .01$.

fixation was stored in the computer. Interobserver reliability assessed on 20% of the sample was high, ranging from .89 to .99 ($M = .95$). The observers did not know how many dots were on each slide for any given trial because their view of the display pairs was occluded by a barrier. Similarly, the experimenter who produced the sound sequences and controlled the slide projectors had no knowledge of the infants' looking behavior because her view was occluded by the apparatus onto which the slides were projected.

Results and Discussion

On average, the infants looked toward at least one member of the display pair on 94% of the trials, and they looked toward both members of the display pair on 66% of the trials. The average duration of attention was 4.78 s out of 10 s. Looking time decreased from an average of 5.32 s on Block 1 trials to 4.15 s on Block 2 trials. These times are comparable with those reported by Starkey et al. (1990).

Each infant's looking times were averaged across all of his or her trials. An analysis of variance (ANOVA) was conducted on these averages,[3] with gender and trial order as between-subjects variables, and trial block (Block 1 vs. Block 2) and display (equivalent vs. nonequivalent) as within-subject variables. This analysis yielded a significant main effect of display, $F(1, 16) = 5.72$, $p < .05$, which reflected longer looking times toward the nonequivalent displays ($M = 2.50$ s, $SD = 0.86$) than toward the equivalent displays ($M = 2.28$ s, $SD = 0.83$). Seventeen of the 20 infants looked longer at the numerically nonequivalent displays. There was also a significant main effect of trial block, $F(1, 16) = 5.72$, $p < .05$, which reflected longer looking times in Block 1 ($M = 5.32$ s, $SD = 2.13$) than in Block 2 ($M = 4.15$ s, $SD = 1.60$). Finally, a significant main effect of gender was found, $F(1, 16) = 5.02$, $p < .05$, which reflected longer average looking times toward both displays by girls ($M = 5.30$ s, $SD = 1.43$) than by boys ($M = 3.89$ s, $SD = 1.54$). No other significant main effects or interactions were found.

Following Starkey et al.'s (1990) data analysis procedures, the proportionate effect of the sounds on looking toward displays of each numerosity also was examined. For each infant, the duration of looking toward displays with two dots when accompanied by two drumbeats was divided by the total duration of looking toward displays with two dots when accompanied by either two or three drumbeats. Likewise, duration of looking toward displays with three dots when accompanied by three drumbeats was divided by looking toward displays with three dots when accompanied by either two or three drumbeats. The resulting proportions were averaged across infants, and then the grand proportions for each numerosity and trial block were compared with the chance value of .50 using two-tailed t tests (see Table 1).

The results of the t tests revealed that across both trial blocks and number of drumbeats, looking time toward the nonequivalent displays was significantly longer, $t(19) = -3.35$, $p < .01$, two-tailed. This held true for both the two- and three-drumbeat sequences, $t(19) = -2.56$, $p < .05$, two-tailed; $t(19) = -2.24$, $p < .05$, two-tailed, respectively. However, when looking times were examined separately for each trial block and numerosity, the only significant effect was found in Block 2 trials for which three was the target numerosity, $t(19) = -2.56$, $p < .05$, two-tailed. This pattern of findings (i.e., significant effects found only in Block 2, the overall analyses, or both) was also reported by Moore et al. (1987) and is similar to the results reported by Starkey et al. (1990), except that the direction of preference is in the opposite direction. There is no obvious explanation for this pattern; however, as Moore et al. suggested, it is possible that infants require a certain amount of exposure to the sound sequences before they can detect the contrast between them.

Starkey et al. (1990) also reported that looking time toward the nonequivalent displays decreased reliably more from Block 1 to Block 2 (0.48 s per trial) than did looking time toward the equivalent displays (0.09 s per trial). However, in the present study this effect was not found. Looking time to both the equivalent and nonequivalent displays decreased from Block 1 to Block 2 at the same rate, 0.64 s per trial and 0.53 s per trial, respectively; $t(19) = 0.35$, *ns*, two-tailed. Starkey et al. further reported that in the first block of trials, infants looked significantly longer toward displays of three objects than toward displays of two objects, regardless of the number of drumbeats that were presented. In Block 1 of the present study, infants showed a similar preference for the three-object display, although this

[3] All results reported are based on analyses of the raw data. However, because the data are proportional, it is possible that the assumptions required for ANOVAs have been violated. Even though an inspection of the data indicated that it was normally distributed, analyses also were conducted on log transformations of the looking times. The results of these analyses were parallel to those found in analyses on the raw data, here and throughout.

preference did not reach significance at $\alpha = .05$, $t(19) = 1.89$, *ns*, two-tailed. No preference for either display was evident in Block 2, $t(19) = 0.07$, *ns*, two-tailed.

Next, we examined patterns of results for individual infants. If, as Starkey et al. (1990) claimed, infants in this procedure are establishing a one–one correspondence between the sounds and items in the visual displays, then this should be reflected in each infant's looking times. We used two statistical tests to determine whether this was the case. First, each infant's average looking times toward the equivalent and nonequivalent displays were compared using a two-tailed t test. This comparison revealed that none of the individual infants showed a significant preference for either display. Next, sign tests were conducted on the direction of preference across each infant's 32 trials. These also revealed no significant difference for any individual infant. Thus, the significant effects observed in the group data were due to a slight ($M = 0.33$ s, $SD = 0.23$) preference for the nonequivalent display shown in the majority of infants (17 out of 20).

The results of the present experiment reveal a weak but significant effect of auditory information on infants' looking behavior. However, contrary to the findings reported by Starkey et al. (1990), infants in the present experiment looked longer toward the display that was not numerically equivalent to the set of sounds. Thus, the present findings are consistent with the results of Moore et al.'s (1987) replication attempt. This is important to note because it strongly indicates that the minor variations in procedure introduced in the present study, such as using dots in the visual displays and randomizing the drumbeat sequences across trials, do not explain the observed preference for the nonequivalent display. Recall that Moore et al. still found a preference for the nonequivalent display even though they did not alter the Starkey et al. procedure in these particular ways. Furthermore, although Starkey et al. contended that the preference for the nonequivalent display reported by Moore et al. was due to a novelty preference set up by allowing some babies to take breaks between trial blocks, the same preference for the nonequivalent display was observed in the present study even though no breaks were given. Thus, there is no clear explanation for the reversal in preference observed across experiments.

Of course, it might be argued that this reversal is inconsequential because a significant preference in either direction demonstrates a sensitivity to numerical information presented in different modalities. However, the view that direction of preference is not meaningful is inconsistent with Starkey et al.'s (1990) use of one-tailed t tests to analyze the proportionate effect of sound on looking behavior. The use of one-tailed tests implies that one direction of preference was predicted and that the direction of preference was indeed viewed as important by these investigators. Furthermore, the fact that the infants' direction of preference shifts between experiments raises questions about uncontrolled variables that may be affecting looking behavior. If such variables were correlated with number, then the observed effects could be related to these variables, rather than numerical information per se. This issue is examined next in Experiment 2.

Experiment 2

Starkey et al. (1990) reported that infants' performance in their study was not affected by the rate or duration of the drumbeat sequences. They based this conclusion on the finding that in two experiments in which all of the drumbeat sequences were presented at a constant rate regardless of numerosity, infants showed the same preference for the matching display as did infants in a third experiment in which all of the drumbeat sequences were presented at a constant duration. However, this procedure does not properly control for the possible effects of temporal information because infants could have used duration information in the first case and rate information in the second. To ensure that infants are not responding to differences in rate, duration, or both, these variables would have to be randomly intermixed within the trials presented to each infant. Experiment 2 introduces this variation in procedure to test whether temporal information formed the basis of preferences shown in previous studies of infants' sensitivity to intermodal numerical correspondences.

Method

Participants. Twenty healthy full-term infants (10 boys and 10 girls) participated in the experiment. They ranged in age from 6 months 6 days to 8 months 26 days ($M = 7$ months 10 days). Seven additional infants were excluded because they either cried or fell asleep.

Stimuli. The stimuli in Experiment 2 were the same as those used in Experiment 1, with one exception. In Experiment 1, the dots had been randomly placed on an imaginary 3×3 matrix. As noted previously, for the displays with two dots, this always resulted in arrangements that would form lines if the dots were connected. For the displays with three dots, random placement almost always resulted in arrangements that would have formed triangles (30 out of 32 trials). In Experiment 2, the arrangements of the three-dot displays were controlled to form lines on half of the trials and triangles on the rest, as in Starkey et al. (1990) and Moore et al. (1987). The three-dot displays with linear arrangements were then randomly intermixed with the displays with triangular arrangements.

Procedure. The procedure was identical to that used in Experiment 1 except for the critical manipulation of randomizing rate and duration information within the trials presented to each infant. This was accomplished in the following way. All of the three-drumbeat sequences were presented at a rate of one beat per second for 3 s as in Experiment 1. However, half of the two-drumbeat sequences were presented on the first and third seconds of the 3-s interval, and half were presented on the second and third seconds of the 3-s interval. Thus, half of the two-drumbeat sequences were equated with the three-drumbeat sequences for duration (i.e., 3 s overall) so that rate varied with numerosity, and half were equated for rate (i.e., one beat per second) so that duration varied with numerosity. These drumbeat sequence types were randomly intermixed within each trial order.

In this experiment, 13 infants completed both trial blocks, receiving a total of 32 trials each. The remaining 7 infants completed 25–31 trials. Five of these infants completed both blocks, but trials were lost because of equipment failure. As before, infants were not given breaks between trial blocks. Interobserver reliability assessed on 20% of the infants was high, ranging from .94 to .99 ($M = .97$).

Results and Discussion

On average, the infants looked toward at least one member of the display pair on 92% of the trials, and they looked at both members of the display pair on 65% of the trials. The average duration of attention was 4.34 s out of 10 s. Looking time decreased from an average of 5.22 s in Block 1 to 3.46 s in Block 2.

An ANOVA was conducted on the infants' average looking times, with gender and trial order as between-subjects variables, and trial block (Block 1 vs. Block 2) and display (equivalent vs. nonequivalent) as within-subject variables. As in Experiment 1, there was a significant main effect of trial block, $F(1, 16) = 13.34, p < .005$, which reflected longer looking times in Block 1 than in Block 2. However, the difference between looking times toward the equivalent and nonequivalent displays was not significant ($p > .25$). No other significant main effects or interactions were found. Also, as in Experiment 1, there was no significant preference shown for either the two- or three-dot display in either trial block, regardless of the number of drumbeats that were played, Block 1: $t(19) = 0.26$, *ns*, two-tailed; Block 2: $t(19) = 0.21$, *ns*, two-tailed.

When the proportionate effect of the sounds on looking toward displays of each numerosity was examined, no significant deviations from chance were observed (see Table 2). Furthermore, there was no difference in the rate of decrease in looking time from Block 1 to Block 2 between the equivalent and nonequivalent displays (0.91 s per trial and 0.86 s per trial, respectively), $t(19) = 0.02$, *ns*, two-tailed. Finally, not one infant showed a significant difference between looking toward the equivalent and nonequivalent displays on the basis of either *t* tests or sign tests performed on each individual infant's data.

In short, when the auditory sequences were presented so that rate and duration varied randomly with numerosity, infants did not show a significant preference for either the numerically equivalent or nonequivalent visual display. This finding contrasts with Starkey et al.'s (1990) claim that infants' responses in this task are based on the number of sounds in each auditory sequence, rather than temporal features of the sets.

General Discussion

In this study, we investigated whether infants could detect the numerical correspondence between a set of sounds and an equivalent set of items in a visual display. Our purpose was to provide a further test of this ability in order to resolve inconsistencies in the literature related to this question. When Starkey et al. (1990) presented infants with pairs of visual displays followed by a set of drumbeats, the infants looked significantly longer toward the equivalent displays. On the basis of this finding, Starkey et al. concluded that infants "represent sets of visible or audible entities in a way that preserves the discreteness of the individual entities" (p. 124) and then compare these representations by performing "computations," such as establishing one–one correspondence. Starkey et al. further proposed that to make these comparisons, infants use either analogical representations of the items in each set or the resulting values of a symbolic tagging process applied to each set. This account implies that infants operate on exact, possibly amodal, representations of number. However, the present results do not support these claims.

First, it should be noted that even if infants can detect the numerical relation between sounds and items in a visual display, this would not necessarily be based on the establishment of an exact one–one correspondence. In Experiment 1, significant effects were found only in the group data; no individual infant showed a significant preference for either display. As Huttenlocher, Jordan, and Levine (1994) have argued, the fact that significant effects in visual habituation experiments are found only by averaging over infants and trials leaves open the possibility that infants operate on approximate, rather than exact, representations of quantity. So, if infants in the intermodal task detect a relation, this relation could be based on approximate quantity (i.e., more sound and more visual items), rather than an exact mapping between entities.

Furthermore, the significant looking preferences in this intermodal task may be based on rate and duration cues rather than numerosity per se. In Experiment 2, when the rate and duration of the auditory sequences was varied within infants, no significant preference was shown for either visual display. This suggests that for auditory sequences, infants might form representations based on temporal features that happen to correlate with numerosity. Although it is unknown from existing studies, this might also be the case for any temporally distributed set regardless of modality (e.g., light flashes). In contrast, visual habituation experiments have consistently shown that infants react to changes in set size despite variations in features that correlate with numerosity, such as line length or density (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981).

Finally, it is possible that there is not a real preference in this intermodal task. That is, significant findings reported in the literature on infants' detection of auditory–visual numerical correspondences might represent only a sample of the studies that have actually been carried out, (i.e., the "file drawer problem"; Rosenthal, 1979). Thus, the published studies could be

Table 2
Looking at Equivalent and Nonequivalent Displays in Experiment 2

	Mean duration of looking (s)		Proportionate effect of sound[a]		
Trial block	Equivalent display	Nonequivalent display	Two-dot display	Three-dot display	Overall
Block 1	2.69	2.61	.49	.49	.51
Block 2	1.83	1.80	.55	.48	.51
Blocks 1 and 2	2.26	2.21	.50	.49	.52

[a] Proportionate effect of sound refers to the percentage of looking time that infants attended to a given display when the equivalent number of drumbeats was played. No proportion differed significantly from chance at $\alpha < .05$.

in the tails of a normal distribution of findings with many unpublished nonsignificant findings in the center. Experiment 2 of the present study is an example of at least one attempt to test this effect that resulted in no significant preference. Similar results may have been found in other laboratories but are not widely known because of a bias to publish only results that reach a certain level of significance (e.g., $\alpha < .05$). Support for the file drawer explanation comes from the fact that the existing studies report small effects in both directions of preference.

Whether or not the reported effects using this procedure are attributable to the use of rate and duration information or to chance, it does not mean that infants lack any numerical competence. As noted previously, there is currently sufficient evidence to suggest that infants are sensitive to the numerical information in visual sets and are capable of making visual–visual comparisons of numerosity. Our findings simply indicate that this sensitivity does not generalize to the present intermodal task.

One reason for this might be that infants cannot represent the numerosity of discrete items within an auditory sequence. Perhaps infants have difficulty performing temporal integration over the sound sequences to group them into sets that can be evaluated numerically. Alternatively, infants may find sounds more difficult to individuate than items in a visual display. Because within mode auditory tasks have not been used to assess infants' numerical discrimination, little is known about their ability to represent the numerosity of auditory sequences.

A second reason that infants might not demonstrate numerical competence on the present task is that they cannot establish numerical correspondence across modalities. Even if it had been shown that infants could represent the numerosity of sets in different modalities, there would be no reason to assume that they could recognize this type of intermodal relation. Recognizing numerical equivalence across modalities involves evaluating an abstract relation between sets. This is quite different from the types of competence tested in previous intermodal experiments, which focus mainly on infants' recognition of objects based on features experienced in different modalities (e.g., Gibson & Walker, 1984; Gottfried, Rose, & Bridger, 1977; Meltzoff & Borton, 1979).

The failure to recognize numerical equivalence for different set types found in the present study has also been observed in studies of numerical competence in preschool children. As noted previously, Mix et al. (1996) found that 3-year-olds performed at chance on a nonverbal auditory–visual numerical matching task, even though they successfully completed a parallel visual–visual matching task. In another investigation using the same procedure, Mix (1995) found a gradual progression from recognition of equivalence between only homogeneous sets of objects presented all together at age 3 years to recognition of numerical equivalence between visual arrays and a variety of sets, such as sequentially presented events, by age 4½ years. These findings indicate that children's ability to recognize numerical equivalence is initially limited to visual sets presented all at once. Existing evidence of numerical competence in infancy suggests that this constraint may be evident from the beginning.

References

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

Gelman, R., & Brenneman, K. (1994). First principles can support both universal and culture-specific learning about number and music. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 369–390). New York: Cambridge University Press.

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

Gibson, E. J., & Walker, A. (1984). Development of knowledge of visual–tactual affordances of substance. *Child Development, 55,* 453–460.

Gottfried, A. W., Rose, S. A., & Bridger, W. H. (1977). Cross-modal transfer in human infants. *Child Development, 48,* 118–123.

Huttenlocher, J., Jordan, N., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General, 123,* 284–296.

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

Mix, K. S. (1995). *Preschoolers' recognition of cardinal equivalence.* Unpublished doctoral dissertation, University of Chicago.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (1996). Do preschool children recognize auditory–visual numerical correspondences? *Child Development, 67,* 1592–1608.

Moore, D., Benenson, J., Reznick, J. S., Peterson, M., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology, 23,* 665–670.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43,* 93–126.

Starkey, P., & Cooper, R. G., Jr. (1980, November 28). Perception of numbers by human infants. *Science, 210,* 1033–1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983, October 14). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179–181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97–128.

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

Wynn, K. (1992). Evidence against empiricist accounts of origins of numerical knowledge. *Mind and Language, 7,* 315–332.

Received December 15, 1995
Revision received July 8, 1996
Accepted July 9, 1996 ■