Do Preschool Children Recognize Auditory-Visual Numerical Correspondences?

Kelly S. Mix, Janellen Huttenlocher, and Susan Cohen Levine

University of Chicago

MIX, KELLY S., HUTTENLOCHER, JANELLEN; and LEVINE, SUSAN COHEN. *Do Preschool Children Recognize Auditory-Visual Numerical Correspondences?* CHILD DEVELOPMENT, 1996, 67, 1592–1608. The present study investigated the ability of 3- and 4-year-old children to perform tasks which require matching sets of sounds to numerically equivalent visual displays. We found that 3-year-olds performed at chance on the auditory-visual matching task, while 4-year-olds performed significantly above chance. There is evidence that mastery of the linguistic counting system is related to success on this task. These findings are unexpected given previous research reporting that 6–8-month-olds can detect the numerical equivalence between a set of sounds and items in a visual display.

Several experiments have tested the ability of infants to discriminate between small sets. It has been shown that after being habituated to a small array of items, infants from birth to 12 months of age will dishabituate when shown an array of a different numerosity (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 did not rely on these factors as the basis for making discriminations. However, it was not clear what particular mechanism was most likely underlying the infants' performance.

Starkey, Spelke, and Gelman (1990) argued that by using only visual stimuli the habituation studies did not rule out the possibility that infants had used "a visual numerosity detection process called subitizing rather than a more central process" (p. 100). They proposed that evidence of the ability to recognize numerical correspondences between more disparate sets of items, such as visual displays and sounds, would ensure that infants' responses were based on the detection of numerical information. They reasoned that this ability would depend on a process involving one-to-one correspondence rather than on visual subitizing, since the latter process could not be applied to temporally distributed sets of sounds.

To test this ability, Starkey et al. (1990) designed a series of experiments in which 6–8-month-olds were presented with both visual and auditory sets. One approach was to show pairs of visual displays with sets of two and 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. In a different test of this matching ability, infants were required to detect the correspondence between a set of sounds and a visual display when the two were not present simultaneously. Infants were shown displays of either two or three objects until looking time had decreased to a set criterion. Then they heard either two or three drumbeats which emanated from behind a black disk that was projected onto a screen. The infants looked significantly longer toward the source of the sounds when the sequences matched the numerosity of the habituation displays.

Starkey et al. (1990) 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

We thank Martha Wagner Alibali, Miriam Bassok, Rebecca Compton, Erin C. Pitts, Elisabeth Hollister Sandberg, and anonymous reviewers for their helpful comments on earlier drafts of this article. Correspondence concerning this article should be addressed to Kelly S Mix, Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637 Electronic mail may be sent via Internet to mix1@midway.uchicago.edu.

[*Child Development*, 1996, **67**, 1592–1608 © 1996 by the Society for Research in Child Development, Inc.
 0009-3920/96/6704-0014$01.00]

of numerical equivalence. They concluded that in order to detect such relations, infants must make use of a process involving both one-to-one correspondence and the abstraction principle (i.e., knowledge that any discrete element, including sounds, can be enumerated) (Gelman & Gallistel, 1978). Further, they proposed that the emergence of these abilities is dependent on neither the acquisition of language nor a culture-specific counting system.

Clearly, this interpretation attributes to infants competence far beyond that shown in previous studies of infant number concepts and, indeed, beyond that shown in young children. It is reasonable to predict that if infants can recognize auditory-visual numerical correspondences, children should succeed on similar tasks. In fact, it has been claimed that acquisition of the conventional number system is guided by the preverbal numerical competencies available in infancy (Gallistel & Gelman, 1992; Gelman, 1991). Thus, if infants have an abstract number concept, this should be evident in early childhood.

The competence tested by Starkey et al. (1990) has not been directly assessed in young children. However, existing studies of counting system acquisition suggest that children are not guided by an abstract number concept. For example, Schaeffer, Eggleston, and Scott (1974) asked children to produce various sets ranging from one to seven items. In one condition, children produced sets of candies and in another condition they produced sets of sounds by tapping on a drum. For less proficient counters, performance in the drum condition was significantly worse than performance in the candy condition. Performance in both conditions improved and began to converge once children had mastered the conventional counting system. In another study, Wynn (1990) asked preschool children to count various sets of items, including physical objects, events, and sounds. For all age groups tested, the percentage of successful counts was lower in the sound condition than for any of the other conditions. In fact, several of the $2\frac{1}{2}$-year-olds appeared bewildered by the conditions involving either events or sounds and failed to count in them even though they had successfully counted physical objects. Finally, Shipley and Shepperson (1990) found that young children demonstrate a clear preference for counting discrete physical objects and have great difficulty overcoming this bias in order to count nonobjects, such as classes or parts of items.

These findings suggest a lack of continuity with the claims made by Starkey et al. (1990). However, because the tasks used in these studies required children to count, it is unclear whether the results reflect a lack of conceptual understanding about number or a lack of experience with the linguistic counting routine. That is, children may fail to count nonobjects simply because they have never seen the count words applied in this situation. It is unknown whether preschool children could recognize auditory-visual numerical equivalence in a task that does not require counting. That is, if children can use a nonverbal method for establishing numerical equivalence between auditory and visual sets, then a nonverbal matching task more similar to the Starkey et al. (1990) procedure would provide a better test.

The present study investigates the ability of preschool children to make judgments of numerical equivalence across modalities using a task that does not require counting. Three experiments were conducted to determine whether preschool children could successfully complete tasks similar to those used in Starkey et al.'s infant studies (1990). All three experiments involved presentation of a target set followed by the child's choosing an equivalent set from between two choice arrays. This approach avoids the possible confounding of counting ability and numerical reasoning ability described above. In Experiment 1, the choice arrays were presented after the target stimuli. In Experiment 2, the choice arrays were presented before the target stimuli and remained visible throughout the trial. In Experiment 3, the choice arrays were presented after the target stimuli, as in Experiment 1, but children were given explicit instructions.

Experiment 1

In this experiment, children's ability to match visual arrays to auditory stimuli was tested when the arrays and the stimuli were presented in succession. This was compared to their ability to match the same visual arrays to visual stimuli. The auditory-visual matching task provides a test of the competence that distinguishes Starkey et al.'s (1990) study from previous work on infant number concepts, namely, the ability to recognize intermodal numerical correspondences. The visual-visual matching task controls for overall task complexity by testing three basic components of the auditory-visual matching task. These include (1) the ability to complete a match-to-sample task based on numerical information, (2) the ability to indicate one's choice by pointing, and (3) the ability to induce the goal of the experimental task through demonstration and feedback in a set of familiarization trials. Two counting tasks were included to determine whether success on either matching task is related to conventional counting ability. The How Many task measures children's ability to recite the count word sequence to enumerate a set of objects, as well as their understanding that the final word used in a count represents the set's cardinal number. The Give-a-Number task further measures children's ability to use the count words to represent the cardinality of a set but has a format that does not require understanding of quantitative terms, such as "many" or "how many," which may be difficult for young children to interpret (Wynn, 1990).

Method

Subjects.—Ninety-six children participated in the experiment. They were divided evenly into two age groups (years-months). 3-year-olds (M = 3-6; range 3-0 to 3-11) and 4-year-olds (M = 4-6; range 4-0 to 4-11).[1] These age groups were chosen in an attempt to test children both before and after mastery of the linguistic counting system. Each group included 24 boys and 24 girls. The children were drawn from preschools that served a predominantly white, middle-class population in the greater Chicago area. All came from homes where English was the primary language.

Materials and procedure.—Each child completed two matching tasks: auditory-visual matching and visual-visual control; and two counting tasks: How Many and Give-a-Number. The auditory-visual matching and visual-visual control tasks were always presented before the counting tasks and were counterbalanced for order of presentation across subjects. The How Many task was presented next, followed by the Give-a-Number task.

For both matching tasks, children were presented with a target set and then they chose from between two arrays of dots the one that matched the target set. Two sets of 5 × 8 inch, unlined, white index cards were used for the arrays on both of the matching tasks. Each card had a horizontal line of ¾-inch black dots, ranging in number from one to five, which ran across its center. The backs of the cards were covered in black poster board so that the dots were visible only when the cards were facing up. One set of cards was drawn so that all of the lines of dots were of equal length. A second set was drawn so that the density of the dots was held equal.[2] Half of each set was mixed with half of the other in a random order, resulting in the formation of two sets for which half of the trials were controlled for line length and half were controlled for density. Sample trials of each type are displayed in Figure 1. Each child was presented with one of the two choice card sets for both matching tasks.

Each set of cards was divided into pairs that included a target numerosity, either two, three, or four, and a foil that was the target numerosity plus or minus one. Within each pair, both cards were controlled for either line length or density. That is, a density-controlled card was never paired with a line length–controlled card. The side on which the target numerosity appeared was counterbalanced across trials, so, for example, half of the target "two's" appeared on the left and half appeared on the right. Each pair constituted one trial.

Auditory-visual matching trials began with the cards placed face down in two stacks approximately 9 inches apart. Then the experimenter presented the target numerosity as series of claps. For example, on a "two trial" the experimenter would clap two times. Next, the first pair of choice cards was immediately flipped over to reveal the arrays of dots and the child indicated his or her choice by pointing. Claps were presented at an even rate of one per second. The experimenter sat directly across from the child and her hands were fully visible throughout presentation of the claps. Unlike the infant experiments which provided only auditory stimuli, this mode of presentation provided children with additional numerical information that was visual-sequential. Furthermore, this avoided potentially distracting procedures such as presenting the sounds from a hidden location or turning a tape recorder on and off.

Visual-visual control trials were completed in a similar manner, except that target

[1] The children were originally divided into four 6-month age groups, however, for the sake of clarity and ease of interpretation, these were collapsed into two age groups. Statistical tests reveal the same pattern of findings whether or not the age groups are collapsed.

[2] Of course, it would be impossible to vary line length and density on cards with only one dot. For these pairs, the single dot always appeared in the center of the card.

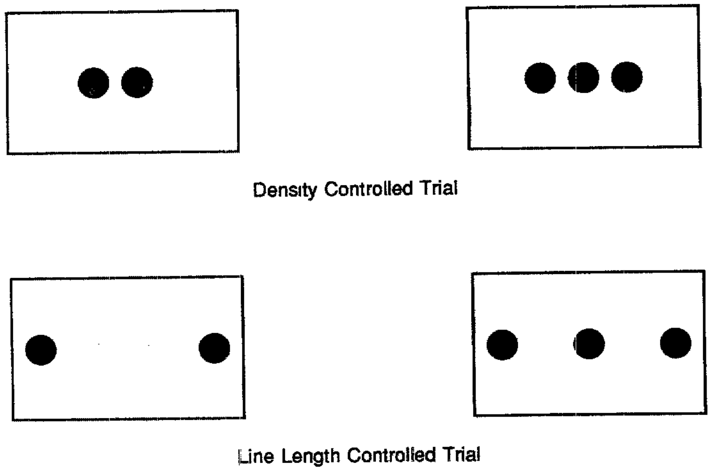

FIG. 1.—Example choice card pairs used for the auditory-visual and visual-visual matching tasks

numerosities were presented as sets of objects. On a blank white card in between the two stacks of choice cards, an experimenter placed the target number of ¾-inch black disks in a horizontal line. On trials where the choice cards were controlled for density, the line of disks was presented as they would appear on a line length–controlled card. Similarly, when the trial had line length–controlled choice cards, the disks were laid out as they would appear on a density-controlled choice card (see Fig. 1). The disks were left in full view of the child for a few seconds and were then covered with a white box. Next, the choice cards were flipped over to reveal the arrays of dots, and the child indicated his or her choice by pointing.

Each matching task was preceded by a brief series of familiarization trials, on the pair of one versus two, which were presented in the following way. First, the experimenter said, "We're going to play a game. I'll show you how it goes." Then, she demonstrated the task by presenting a target set and pointing to the correct card while they were still facing up. The child was told, "Now it's your turn," and received two practice trials of this type. Next, the task was demonstrated with the choice cards facing down, as is the case on test trials, and the child received two more practice trials of this type. These were followed immediately by the 12 test trials, in which the cards were face down. Most children readily grasped the task and pointed without any further instruction. However, eight children (four from each age group) who initially pointed to both choice cards during familiarization were asked to point to the card that was the same. Following the verbal prompt, all eight children responded appropriately by pointing to only one card. When a correct response was given during the practice trials, the child was praised and told, "That's right!" If an incorrect response was given, the child was shown the correct answer and told, "It was this card." No feedback of any kind was given during test trials.

The How Many and Give-a-Number tasks were given to measure mastery of the conventional counting system. In the How Many task, children were given a long piece of corrugated cardboard with 10 disks glued in a line down the middle. They were asked to count the disks aloud, and then they were asked to tell how many disks there were. Ten disks were chosen in order to provide a sufficiently challenging range. In the Give-a-Number task, children were given 15 disks and asked to place a certain number of them on a blank index card. Each of the numbers from one to six was requested in one of two fixed random orders. Once the child had responded to each request, the disks were returned so that the pile of 15 disks remained constant. The range of numerosities requested is based on Wynn's (1990) procedure.

Results

For every child, each correct match was given one point, resulting in a total possible score of 12 for each matching task. A total score of six is predicted by chance because on any given trial there is a .50 probability of answering correctly by guessing. The matching task performance for children in

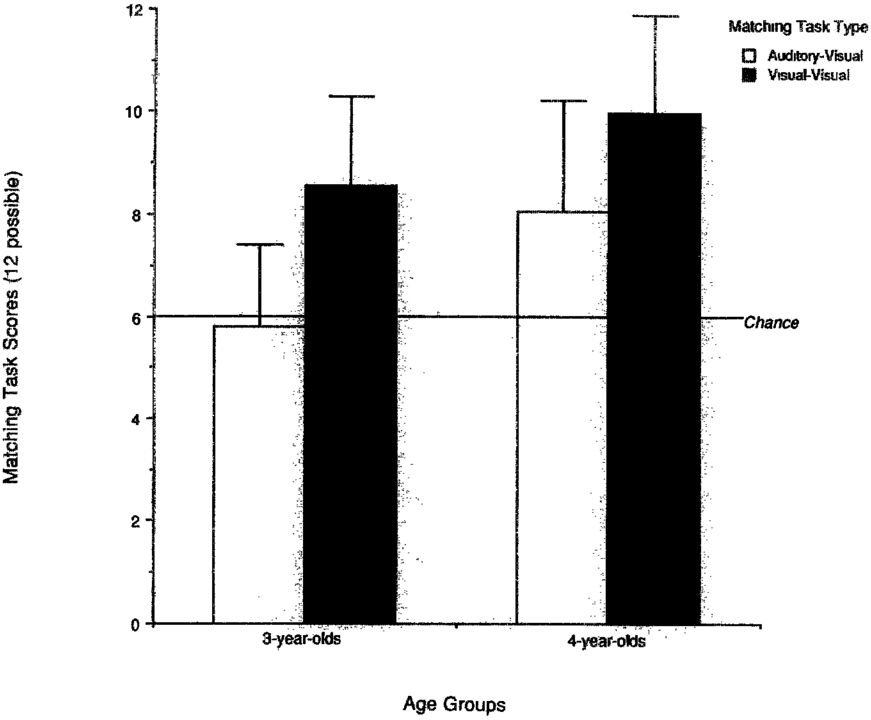

FIG. 2.—Matching task performance as a function of age group and matching task type for Experiment 1

each age group is presented in Figure 2. Two-tailed t tests were used to compare the mean scores for children in each age group with the score predicted by chance (i.e., 6 out of 12). These comparisons revealed that 3-year-olds performed at chance on the auditory-visual matching task, $t(47) = 0.82$, N.S., whereas the 4-year-olds performed significantly above chance, $t(47) = 6.59, p < .0005$. In contrast, both groups performed significantly above chance on the visual-visual control task: 3-year-olds, $t(47) = 10.29, p < .0005$, 4-year-olds, $t(47) = 14.29, p < .0005$.

A similar pattern of results emerged when the data were analyzed according to the number of children whose performance exceeded a criterion of nine out of 12 correct matches on each matching task ($p < .06$) (see Table 1). The number of children expected to pass this criterion can be determined by multiplying the probabilities associated with the criterion with the number of children in the sample. Hence, based on random guessing, 2.88 of the 48 children in each age group would be expected to pass this criterion and 45.12 would not. As shown in Table 1, this result was obtained for 3-year-olds' performance on the auditory-visual task, $\chi^2(1) = .01$, N.S.; however, 25 of the 48 3-year-olds passed the criterion for the visual-visual matching task, $\chi^2(1) = 180.74, p < .0005$. In contrast, the number of 4-year-olds passing this criterion was significantly above chance for both the auditory-visual task, $\chi^2(1) = 121.28, p < .0005$, and the visual-visual control task, $\chi^2(1) = 430.03, p < .0005$.

An analysis of variance was conducted on the children's matching task scores with task type as a within-subject factor and age, gender, choice card set, and order of presentation as between-subjects factors. A significant main effect of task type was revealed, $F(1, 91) = 107.06, p < .0001$, which reflected higher scores on the visual-visual matching task than on the auditory-visual matching task. There was also a significant main effect of age, $F(1, 91) = 35.19, p < .0001$, which reflected better performance by the 4-year-olds ($M = 9.02$) than by the 3-year-olds ($M = 7.19$). Although no main effect of order of presentation was found, there was a significant interaction between task type and order, $F(1, 91) = 5.77, p < .05$. However, pairwise comparisons (Scheffé S, $p < .025$ to control for multiple comparisons) revealed that the differences between performance for each task when it was presented first ver-

TABLE 1

NUMBER OF CHILDREN PERFORMING AT OR ABOVE CRITERION FOR EXPERIMENT 1

Age Group and Task	Number (Proportion)	Average Age (Years-Months)
3-year-olds (n = 48):		
Auditory-visual	3 (.06)	3–9
Visual-visual	25 (.52)	3–6
4-year-olds (n = 48):		
Auditory-visual	21 (.44)	4–7
Visual-visual	37 (.77)	4–6

NOTE.—The criterion used is 75% (9 out of 12) correct matches (p < .06).

sus second did not reach significance (auditory-visual: presented first, M = 6.63, presented second, M = 7.25; visual-visual, presented first, M = 9.04, presented second, M = 9.50). Moreover, performance on the visual-visual matching task was significantly better than on the auditory-visual matching task for both orders of presentation. No other significant main effects or interactions were found. It should be noted that because the data in this experiment are proportional, the assumption of homogeneity of variance could be violated. However, a parallel analysis using arcsin transformations of children's matching task scores revealed the same pattern of results, confirming the robustness of the results reported here.

A second analysis of variance was carried out to examine differences in performance based on the size of the target sets. Recall that four trials for each of the three target set sizes were presented; two, three, and four. A total score of two is predicted by chance because on any given trial there is a .50 probability of answering correctly by guessing. These scores were used in an analysis of variance with age as the between-subjects factor.

The ANOVA revealed a significant main effect of target set size, $F(1, 188) = 23.55$, $p < .0001$, and a significant task type × set size interaction, $F(1, 188) = 4.53$, $p < .05$. These effects were modified by a significant three-way interaction between task type, set size, and age group, $F(1, 94) = 3.23$, $p < .05$. (A parallel analysis using arcsin transformations of children's matching task scores revealed the same pattern of results.) An examination of Figure 3 suggests that the interaction is due to the contrast between the 4-year-olds' performance, which showed a sharp decline at set size four on both matching tasks, and the 3-year-olds' performance, which showed no effect of set size on the auditory-visual matching task and a moderate effect of set size for the visual-visual control task. Pairwise comparisons (Scheffé S, $p < .01$ to control for multiple comparisons) of set size for each task and age group confirm that this was the case. On the auditory-visual matching task, 4-year-olds' scores were significantly worse for set size four than for set size two. On the visual-visual control task, children in this age group performed significantly worse for set size four than for either set size two or set size three. In neither case did performance on set size two differ significantly from that on set size three. Similarly, 3-year-olds' visual-visual control task scores for set size four differed significantly from set size two, but scores for set sizes two and three did not differ significantly. In contrast, 3-year-olds' performance on the auditory-visual matching task did not vary significantly based on set size.

It is possible that the ability to recognize auditory-visual numerical matches is related to mastery of the conventional counting system. For example, children who were successful on auditory-visual matching may have relied on their counting skills to meet the demands of this task. If so, one would expect to find that level of conventional counting ability would be related to performance on the auditory-visual matching task scores. The following analyses investigated this question by considering performance on the How Many and Give-a-Number tasks in relation to matching task scores.

For the How Many task, children were first asked to count 10 disks that were glued in a row, and then they were asked to tell how many disks there were. Children re-

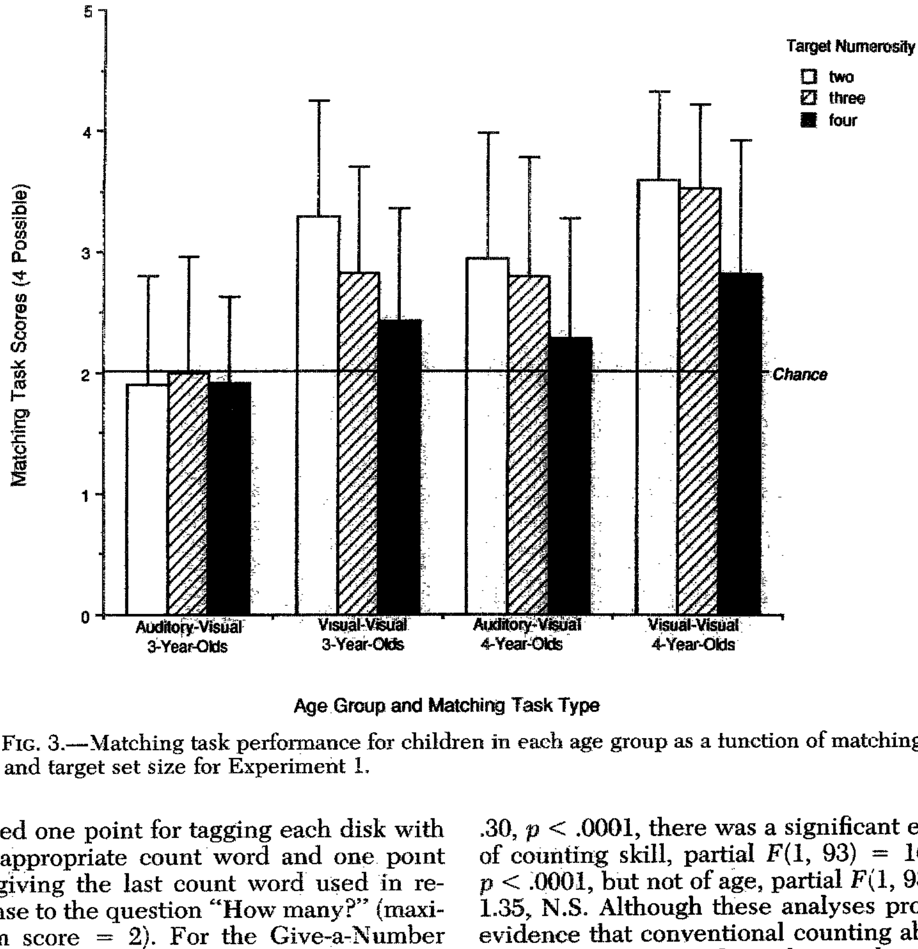

FIG. 3.—Matching task performance for children in each age group as a function of matching task type and target set size for Experiment 1.

ceived one point for tagging each disk with the appropriate count word and one point for giving the last count word used in response to the question "How many?" (maximum score = 2). For the Give-a-Number task, children were asked to produce a set of disks for each numerosity, one through six. Each child received one point for every correct response for each of the numerosities requested (maximum score = 6). All points earned on the How Many and Give-a-Number tasks were pooled together to yield a counting skill score out of eight possible points for each child. Performance on the How Many and Give-a-Number tasks was significantly correlated, $r(95) = .65$, $p < .0001$, providing convergent evidence that a common underlying skill was tapped by both tasks.

The pooled counting scores and the children's ages in months were used as predictors in a multiple linear regression on the children's auditory-visual matching task scores, $r^2(95) = .40$, $p < .0001$. There were significant effects of both counting skill, partial $F(1, 93) = 13.88$, $p < .0005$, and age, partial $F(1, 93) = 9.50$, $p < .005$. In a parallel analysis for visual-visual control task performance using the same predictors, $r^2(95) = .30$, $p < .0001$, there was a significant effect of counting skill, partial $F(1, 93) = 16.11$, $p < .0001$, but not of age, partial $F(1, 93) = 1.35$, N.S. Although these analyses provide evidence that conventional counting ability is related to improved matching task performance, it should be noted that age also is significantly correlated with counting skill, $r(95) = .62$, $p < .0001$. However, even though age and counting skill are highly correlated, there is still a significant effect of counting skill on both tasks.

Furthermore, there is evidence that only proficient counters were successful on the auditory-visual matching task. Children were divided into two groups based on their performance on the counting tasks. Children were coded as being more proficient counters if they obtained at least seven out of eight possible points when performance on the How Many and Give-a-Number tasks was pooled. Children who obtained combined scores of six or less were coded as less proficient counters. Using these criteria, 36 children were identified as more proficient counters (mean age = 4-5; SD = 5.86 months), and the remaining 60 children were identified as less proficient counters (mean age = 4-0; SD = 5.77 months). The matching task performance for both groups, displayed in Figure 4, reveals that more proficient counters performed significantly above chance on the auditory-visual matching task, $t(35) = 8.78$, $p < .001$. In contrast, the mean auditory-visual matching score for children who were less proficient counters was not significantly different from chance, $t(59) = 0.63$, N.S. Mean scores on the visual-visual matching task were significantly above chance for both groups: more proficient, $t(35) = 16.31$, $p < .001$, less proficient, $t(59) = 10.77$, $p < .001$, although somewhat higher for more proficient counters. Thus, even though conventional counting knowledge is related to improved performance on both matching tasks, it may be a prerequisite for success on the auditory-visual matching task.

Based on these results, a second criterion was used to determine whether even minimal counting proficiency (i.e., encompassing the highest numerosity tested) would be sufficient for performing the auditory-visual task. This criterion defined minimal counting proficiency as being able to count at least four items, give up to at least four items, and give the last count word used in response to the question how many (maximum score = 6). Using this criterion, the same pattern of results emerged. Children with minimal counting proficiency ($n = 44$, mean age = 4-3; SD = 6.47 months) were still able to perform the auditory-visual task above chance, $t(43) = 6.17$, $p < .001$. In contrast, those with less than minimal counting proficiency ($n = 52$, mean age = 3-8; SD = 5.64 months) were at chance on the auditory-visual task, $t(51) = 0.08$, N.S. However, children in both groups, even those with less than minimal counting proficiency, could perform the visual-visual matching task above chance: minimal proficiency, $t(43) = 19.12$, $p < .001$, less than minimal proficiency, $t(51) = 9.28$, $p < .001$.

Discussion

The results of Experiment 1 clearly show that the ability to match accurately the number of sounds with an equivalent visual display emerges between 3 and 4 years of age. These findings reveal a gap between the level of understanding attributed to infants by Starkey et al. (1990) and that shown by preschoolers in the present study. If 3-year-olds possess any ability that would allow them to detect numerical correspondences between auditory and visual stimuli, then it is surprising that their scores on the

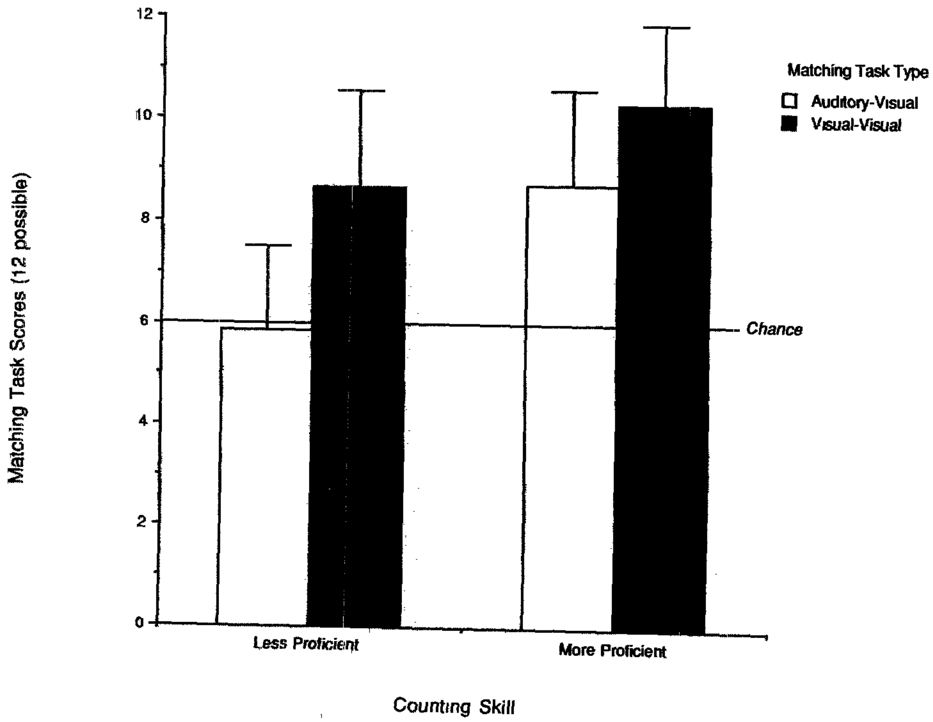

FIG. 4.—Matching task performance as a function of counting proficiency and matching task type for Experiment 1.

present task were not even slightly above chance. Analyses of performance by set size indicated that 3-year-olds performed at chance on the auditory-visual matching task even for the lowest numerosity tested. In contrast, 4-year-olds performed significantly above chance on the auditory-visual matching task, demonstrating that success is attained with development. Children in both age groups performed significantly above chance on the visual-visual control task. This indicates that it was not demands inherent in the experimental procedure that prevented the 3-year-olds from making correct auditory-visual matches. Instead, there appear to be processing demands unique to the auditory-visual matching task that presented a challenge to children in this age group. This could be due to either differences between the visual and auditory modalities or the contrast between the simultaneous presentation of items in a visual array versus the sequential presentation of a set of sounds.

Analyses of performance on conventional counting tasks indicate that mastery of the linguistic counting system is related to success on both matching tasks. However, in contrast to the visual-visual control task, on the auditory-visual matching task it was only children who demonstrated linguistic counting proficiency who were able to perform at a level significantly above chance. These results imply that acquisition of the linguistic counting system precedes the ability to detect numerical correspondences between auditory and visual stimuli.

Experiment 2

In Experiment 1, children were required to represent the total numerosity of a target set in memory and then make the correct numerical match to a visual array. It seemed possible that performance might improve with a different procedure which would facilitate the use of a one-to-one correspondence strategy similar to that which Starkey et al. (1990) proposed to account for the infants' success on their tasks. In Experiment 2, the visual displays and the auditory stimuli were presented simultaneously, thereby allowing matches to be made by mentally tagging each dot as the sounds were presented.

Method

Subjects.—Forty-eight children participated in the experiment. They were divided evenly into three age groups (years-months): 3½-year-olds (M = 3-9; range 3-6 to 3-11), 4-year-olds (M = 4-2; range 4-0 to 4-5), and 4½-year-olds (M = 4-9; range 4-6 to 4-11). Each age group included eight boys and eight girls. Younger 3-year-olds were excluded from this experiment because initial findings indicated that their results would not differ from those of the 3½-year-olds. The children were drawn from preschools that served a predominantly white, middle-class population in the greater Chicago area. None had participated in Experiment 1, although some came from the same preschools as children who had. All came from homes where English was the primary language.

Materials and procedure.—The procedure and materials were identical to those used in Experiment 1, except that on the auditory-visual and visual-visual matching tasks the choice cards were facing up during presentation of either the disks or the claps. For example, an auditory-visual trial would begin when the two choice cards were turned up to reveal the dots. Next, the sequence of claps began. When the claps ended the child made his or her choice. The experimenter indicated that the clapping sequence had ended by folding her hands and looking at the child. Visual-visual trials proceeded in a similar manner. After the choice cards were turned face up, a set of disks was laid out and left in full view of the child for a few seconds. The disks were then covered, after which the child indicated his or her choice by pointing. During the familiarization trials, one 4-year-old did not respond at all. She was prompted to point to the card that was the same and responded by pointing on all trials thereafter.

Results

As in Experiment 1, every correct match received one point, resulting in a total possible score of 12 for each task. A score of six is predicted by chance because on any given trial there is a .50 probability of answering correctly by guessing. The matching task performance for children in each age group is presented in Figure 5. Two-tailed t tests comparing the mean scores for children in each age group with the score predicted by chance (i.e., 6 out of 12) revealed that only the 4½–5-year-old age group performed significantly above chance on the auditory-visual matching task: 3½-year-olds, $t(15)$ = 0.50, N.S., 4-year-olds, $t(15)$ = 0.42, N.S., 4½-year-olds, $t(15) = 2.78$, $p < .01$. However, performance on the visual-visual control task was significantly above chance for all three age groups: 3½-year-olds, $t(15)$ = 5.23, $p < .0005$, 4-year-olds, $t(15) = 7.30$, p

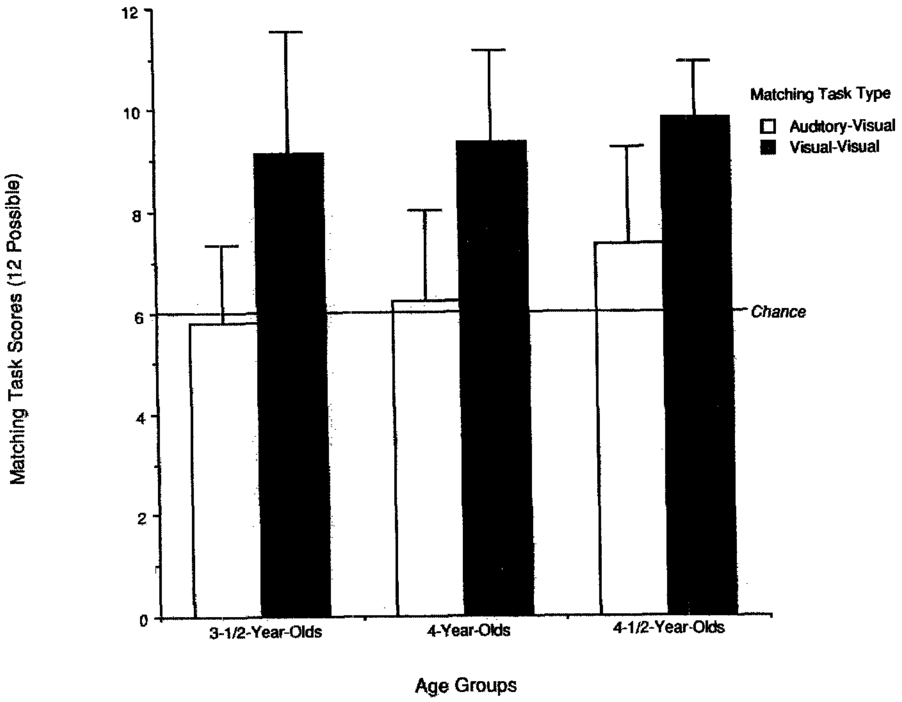

FIG. 5 —Matching task performance as a function of age group and matching task type for Experiment 2.

$< .0005$, 4½-year-olds, $t(15) = 13.34$, $p < .0005$.

A similar pattern of results emerged when the data were analyzed according to the number of children whose performance exceeded a criterion of nine out of 12 correct matches on each matching task ($p < .06$). On the auditory-visual matching task, the number of children passing this criterion exceeded the expected value (i.e., one child) for only the 4½-year-old group: 3½-year-olds, $\chi^2(1) = .002$, N.S., 4-year-olds, $\chi^2(1) = 1.20$, N.S., 4½-year-olds, $\chi^2(1) = 28.15$, $p < .0005$. In contrast, the number of children passing this criterion on the visual-visual control task was significantly greater than the expected value for children in all age groups: 3½-year-olds, $\chi^2(1) = 111.7$, $p < .0005$, 4-year-olds, $\chi^2(1) = 111.7$, $p < .0005$, 4½-year-olds, $\chi^2(1) = 218.44$, $p < .0005$ (see Table 2).

TABLE 2

NUMBER OF CHILDREN PERFORMING AT OR ABOVE CRITERION FOR EXPERIMENT 2

Age Group and Task	Number (Proportion)	Average Age (Years-Months)
3½-year-olds (n = 16):		
Auditory-visual	1 (.06)	3–11
Visual-visual	11 (.69)	3–10
4-year-olds (n = 16):		
Auditory-visual	2 (.13)	4–2
Visual-visual	11 (.69)	4–3
4½-year-olds (n = 16):		
Auditory-visual	6 (.38)	4–10
Visual-visual	15 (.94)	4–9

NOTE.—The criterion used is 75% (9 out of 12) correct matches ($p < .06$)

An analysis of variance on the children's matching task scores with task type as a within-subject factor and age group, gender, choice card set, and order of presentation as between-subjects factors revealed a significant main effect of task type, $F(1, 42) = 72.42$, $p < .0001$, such that scores for the visual-visual matching task were higher than those for the auditory-visual matching task. No other significant main effects or interactions were found.

A second analysis was conducted to examine the effects of target set size. As in Experiment 1, each child's performance was coded separately for each target set size, resulting in a total possible score of four for each numerosity tested on each matching task. (A total score of two is predicted by chance because on any given trial there is a .50 probability of answering correctly by guessing.) These scores were used in an analysis of variance with age as the between-subjects factor. A significant main effect of target set size, $F(1, 90) = 10.06$, $p < .0001$, was qualified by a significant task type × set size interaction, $F(1, 90) = 19.55$, $p < .0001$. An examination of Figure 6 suggests that this interaction is due to the contrast between a steady decline in performance across set size on the visual-visual control task and stable performance at chance across set size on the auditory-visual matching task. Pairwise comparisons (Scheffé S, $p < .025$ to control for multiple comparisons) confirm that the scores for each target set size do not differ significantly from one another on the auditory-visual matching task. On the visual-visual control task, performance on trials for which four was the target set size was significantly worse than for set sizes two or three. The difference between performance on set size two and set size three also approached significance ($p = .026$). Further, an examination of the difference between performance on each matching task within set size (Scheffé S, $p < .01$ to control for multiple comparisons) revealed that this difference was significant for set sizes two and three, but not for set size four. Thus, the decline in performance across set size on the visual-visual control task eliminated the visual-visual performance advantage at set size four. Unlike Experiment 1, there were no significant interactions involving set size and age. Parallel analyses using arcsin transformations of children's matching task scores revealed the same pattern of results as those using the raw data.

Analyses of the relation between counting skill and matching task performance parallel to those used in Experiment 1 were carried out. Children's counting skill was coded following the procedure established in Experiment 1, and their pooled counting scores and ages in months were used as factors in a multiple linear regression on the children's auditory-visual matching task scores, $r^2(47) = .27$, $p < .001$. The results revealed significant effects of both counting skill, partial $F(1, 45) = 6.19$, $p < .02$, and age, partial $F(1, 45) = 5.43$, $p < .05$. In contrast, a parallel regression analysis for visual-visual control task performance did not reach significance, $r^2(47) = .08$, N.S., and there were no significant effects found for either counting skill, partial $F(1, 45) = 1.34$, N.S., or age, partial $F(1, 45) = 1.33$, N.S. As before, it is important to note that age is significantly correlated with counting skill, $r^2(47) = .30$, $p < .05$; however, level of counting skill still has a significant effect on auditory-visual matching task performance.

As in Experiment 1, children were coded as minimally proficient or less than minimally proficient counters based on their ability to count up to four, give at least four items, and give the last count word used in response to the question "How many?" Using these criteria, 21 of the 48 children who participated in Experiment 2 attained minimally proficient counter status (mean age = 4-4; SD = 5.03 months) and 27 did not (mean age = 4-3; SD = 5.16 months). Matching task performance for these two groups is displayed graphically in Figure 7. Minimally proficient counters performed significantly above chance on the auditory-visual matching task, $t(20) = 2.36$, $p < .05$. In contrast, the mean auditory-visual matching score for children who were less than minimally proficient counters was not significantly different from chance, $t(26) = 0.13$, N.S. Average scores on the visual-visual matching task were significantly above chance for both groups: minimally proficient counters, $t(20) = 7.54$, $p < .001$, less than minimally proficient counters, $t(26) = 11.15$, $p < .001$, although slightly higher for the more proficient group. Thus, as in Experiment 1, there is evidence that even though conventional counting knowledge is related to improved performance on both matching tasks, it may be a prerequisite for success on the auditory-visual matching task.

Discussion

The results of Experiment 2 indicate that, as in Experiment 1, there were fewer correct auditory-visual than visual-visual matches. However, unlike Experiment 1,

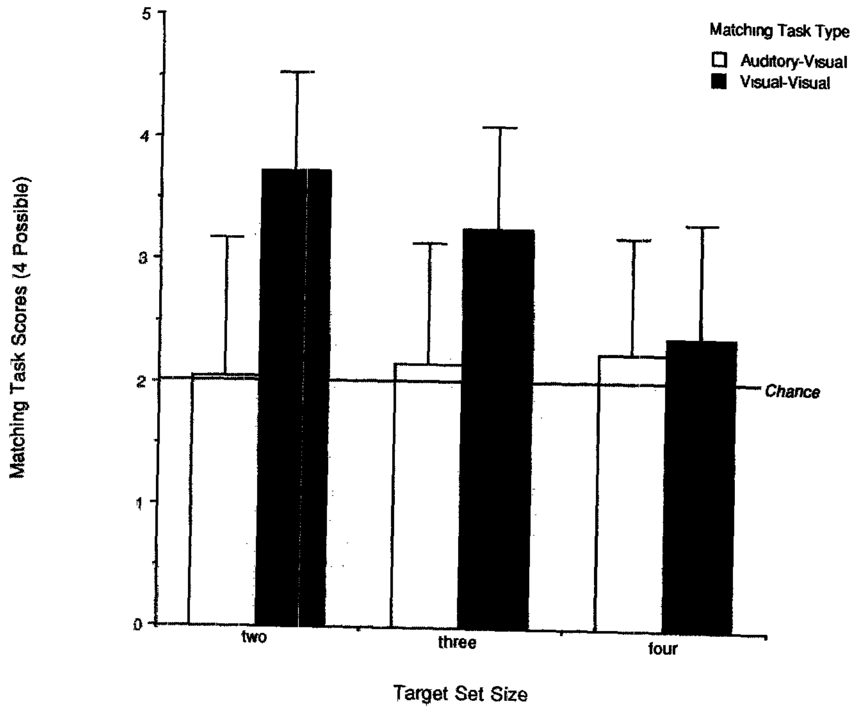

FIG. 6.—Matching task performance for children in each age group as a function of matching task type and target set size for Experiment 2.

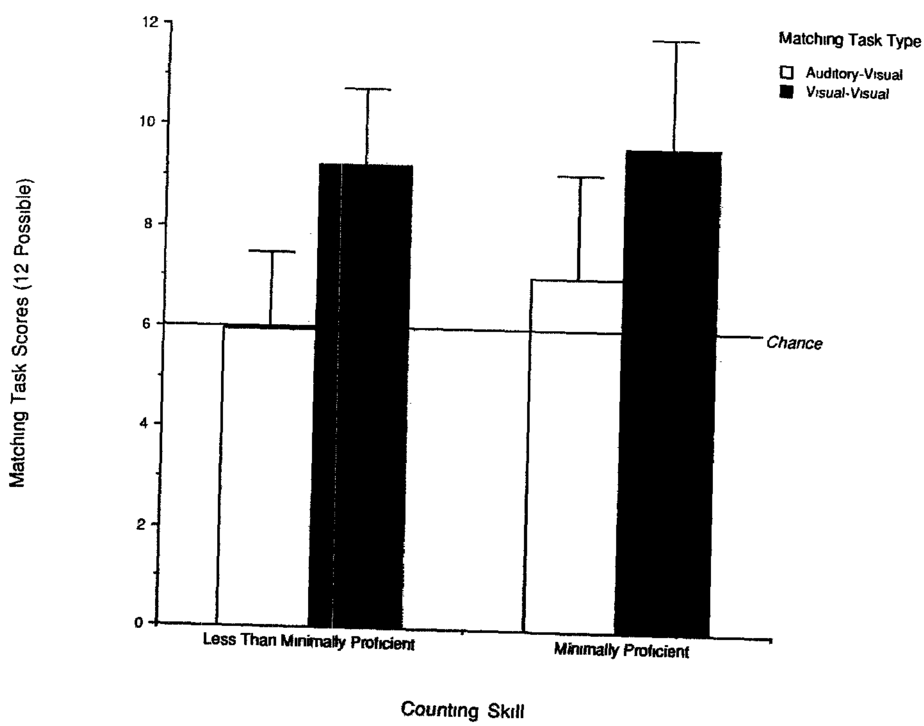

FIG. 7.—Matching task performance as a function of counting proficiency and matching task type for Experiment 2.

there was no significant main effect of age. Even though the mean scores on each task improved steadily with development, the differences in performance among the three age groups were not statistically significant. This may be due to either the greater statistical power provided by the larger sample size in Experiment 1 or because the task used in Experiment 2 was sufficiently difficult that a floor effect was observed for children in age range tested.

In Experiment 2, the target sets and response arrays were presented simultaneously because it seemed possible that this might improve performance by facilitating the use of one-to-one correspondence. This was a reasonable prediction, inasmuch as tagging items in a one-to-one fashion was well within the children's range of ability. When children were given credit for tagging each item on the How Many task, regardless of whether the tags were the appropriate count words, and allowing for one error of either skipping or double-counting an item (e.g., Wynn, 1990), it was found that 81% of the children successfully applied the one-to-one principle to tag individual items. This is consistent with previous evidence that preschool children are quite accurate in applying the one-to-one principle to count small sets (Gelman & Gallistel, 1978; Wynn, 1990). Furthermore, it demonstrates an impressive tagging ability on the part of the children who participated in the current study, given the large set size of 10 items used to test them.

However, the present results indicate that children were not aided by the change from sequential to simultaneous presentation of the target sets and response arrays. In fact, 3½- and 4-year-olds in Experiment 2 performed at chance on the auditory-visual matching task using the simultaneous procedure. This suggests that children may not apply one-to-one correspondence on the auditory-visual task, or, if they do, this strategy may not be particularly effective. For example, on the present task, a child would need to tag mentally each dot in the visual array as the sounds were presented. If the child happened to be looking at the nonmatching array while doing this tagging, he or she would have to deduce that the other card was the correct match by default.

Experiment 3

In Experiments 1 and 2, children were required to induce the goal of the task through demonstration and feedback in order to respond appropriately. Although children's performance on the visual-visual control task indicated that they were able to do so, it is possible that performance on the auditory-visual matching task would improve if explicit instructions were given. If the processing demands of the auditory-visual matching task are greater, then explicit instructions might be required for children to understand the task, whereas explicit instructions would not be required for success on the simpler visual-visual control task. In Experiment 3, children were presented with the same tasks as in Experiment 1, but they were given explicit instructions in addition to demonstration and feedback.

Method

Subjects.—Twenty-four 3-year-olds (M = 3-6; range 3-0 to 3-11) participated in the experiment. There was an equal number of boys and girls. The children were drawn from preschools that served a predominantly white, middle-class population in the greater Chicago area. None had participated in Experiment 1 or Experiment 2, although most came from the same preschools as children who had. All came from homes where English was the primary language.

Materials and procedure.—The procedure and materials were identical to those used in Experiment 1, except that the experimenter introduced the task by saying, "I'm going to clap (or lay out some disks) and you're going to point to the card with the same number. I'll show you how." Then the experimenter demonstrated the task as in Experiments 1 and 2. On the practice trials, children were told, "Now I'll clap (or lay out some disks) and you point to the card with the same number." If children pointed to the correct array they were told, "Yes! That's the card with the same number." If they pointed to the incorrect array, the experimenter pointed to the correct array and said, "*This* card has the same number." After six test trials, children were reminded again to point to the card with the same number.

Because no effect of the choice card set was found in Experiments 1 and 2, only one set (Form 1) was presented in Experiment 3. As before, the order of presentation for the auditory-visual matching and visual-visual control tasks was counterbalanced across subjects. The How Many and Give-a-Number tasks were presented after the matching tasks.

Results

Children in Experiment 3 obtained a mean score of 5.70 on the auditory-visual matching task, and 8.29 on the visual-visual control task. Two-tailed t tests comparing these scores with the score predicted by chance (i.e., 6 out of 12) revealed that children performed at chance on the auditory-visual matching task, $t(23) = 1.49$, N.S. In contrast, performance on the visual-visual control task was significantly above chance, $t(23) = 6.21$, $p < .001$. A similar pattern of results emerged when the data were analyzed according to the number of children whose performance exceeded a criterion of nine out of 12 correct matches on each matching task ($p < .06$). The expected number of children passed this criterion for the auditory-visual task, number passing = 1, $\chi^2(1) = .14$, N.S., but significantly more children passed the criterion for the visual-visual control task, number passing = 12, $\chi^2(1) = 82.38$, $p < .0005$. An analysis of variance on the children's auditory-visual and visual-visual matching task scores with gender and order of presentation as between-subjects factors revealed a significant main effect of task type, $F(1, 21) = 46.18$, $p < .0001$, such that scores for the visual-visual matching task were higher than those for the auditory-visual matching task.

A second analysis was conducted to examine the effects of target set size. As before, every child's performance was coded separately for each target set size, resulting in a total possible score of four for each numerosity tested on each matching task. (A total score of two is predicted by chance because on any given trial there is a .50 probability of answering correctly by guessing.) These scores were used in an analysis of variance with task type as the within-subject factor. A significant main effect of target set size, $F(1, 46) = 3.78$, $p < .05$, was qualified by a significant task type × set size interaction, $F(1, 46) = 8.56$, $p < .001$. This interaction was due to the contrast between a decline in performance across set size for the visual-visual control task, and stable chance level performance across set size for the auditory-visual matching task. Pairwise comparisons (Scheffé S, $p < .025$ to control for multiple comparisons) confirm that, on the visual-visual control task, scores for set size four ($M = 2.38$) were significantly lower than those for set size two ($M = 3.73$), but not for set size three ($M = 3.25$, $p = .036$). Scores for set size two were not significantly different from scores for set size three. However, scores on the auditory-visual matching task did not differ significantly from one another based on target set size (set size two: $M = 2.94$; set size three: $M = 2.15$; set size four: $M = 2.25$). Further, an examination of the difference between performance on each matching task within set size (Scheffé S, $p < .01$ to control for multiple comparisons) revealed that this difference was significant for set sizes two and three, but not for set size four. Thus, the decline in performance across set size on the visual-visual control task eliminates the visual-visual performance advantage at set size four. Parallel analyses of variance using arcsin transformations of children's matching task scores revealed the same pattern of results as those reported above.

Analyses of the relation between counting skill and matching task performance parallel to those used in Experiment 1 were carried out. Children's counting skill was coded following the procedure established in Experiment 1, and both these scores and the children's ages in months were used as factors in a multiple linear regression on the children's auditory-visual matching task scores, $r^2(23) = .18$, N.S. The results revealed a significant effect of counting skill, partial $F(1, 21) = 4.66$, $p < .05$, but not age, partial $F(1, 21) = 0.70$, N.S. In a multiple linear regression on visual-visual control task scores, $r^2(23) = .28$, $p < .05$, counting skill was found to be marginally significant, partial $F(1, 21) = 4.11$, $p < .06$, but age was not, partial $F(1, 21) = 1.20$, N.S. Although only 3-year-olds participated in Experiment 3, and the small age variance limits the statistical effects of age, age and counting skill were marginally correlated, $r(23) = .37$, $p < .10$, indicating that older 3-year-olds had better counting ability than younger 3-year-olds. The significant effect of counting ability on auditory-visual matching task performance, even when counting skill and age are not highly correlated, suggests that knowledge of the conventional counting system makes an important contribution to success on the auditory-visual task.[3]

Discussion

The results of Experiment 3 indicate that providing explicit instructions did not lead to improved performance. As before, 3-year-old children performed at chance on the auditory-visual matching task and sig-

[3] The minimally proficient counter analysis was not carried out in Experiment 3 because only four of the 24 children met the criterion for minimally proficient counter status established in Experiment 1. Since this made the two groups to be compared extremely unbalanced, this analysis would not be reliable.

nificantly above chance on the visual-visual control task. This finding is consistent with other indications that children fail to recognize auditory-visual numerical correspondences even though they understand the basic requirements of the matching task. For example, 3-year-olds' above chance performance on the visual-visual control task provides strong evidence that they were capable of completing a match-to-sample task based on numerical information. Nonetheless, children in this age group consistently performed at chance on the auditory-visual matching task.

It appears that demands unique to the auditory-visual matching task, and not an inability to grasp the general requirements of the task itself, explains the 3-year-olds' failure. Of course, an experiment like this could never rule out the possibility that different instructions might be more effective. However, it is not clear what would be included in such instructions.

General Discussion

Starkey et al. (1990) reported that infants can recognize numerical equivalence across modalities. They posited the use of a process involving both one-to-one correspondence and the abstraction principle and argued that the emergence of this ability is not dependent on the acquisition of a conventional counting system. Based on these claims, one would expect preschool children to show a similar competence. However, the present experiments provide no evidence that 3-year-olds can detect the numerical correspondence between auditory and visual stimuli. This is true even though they successfully completed a control task that required visual to visual matching. In contrast, 4-year-olds performed significantly above chance on both tasks, demonstrating that success on the auditory-visual matching task is attained with development. The counting task analyses suggest that success on the auditory-visual matching task may be related to mastery of the linguistic counting system. The present findings are difficult to reconcile with claims that infants recognize intermodal numerical correspondences and that acquisition of the conventional number system is guided by the preverbal numerical competencies available in infancy (Gallistel & Gelman, 1992; Gelman, 1991).

There are several possible explanations for the lack of continuity between Starkey et al.'s (1990) claims and the present results. One is that the discrepancy is due to the contrast between the use of a passive looking time response in the infant study and an active choice response in the preschool study. Perhaps the sensitivity demonstrated by infants is not available to conscious inspection or application, whereas pointing to a card requires a conscious choice. Gelman and Brenneman (1994) have argued that while infants may possess innate knowledge of numerical principles, they do not have access to this knowledge. Instead, the principles are represented implicitly within the structure of the information processing mechanisms that assimilate and direct the infants' actions. Yet, if the reason 3-year-olds failed to detect intermodal numerical correspondences is because the preschool task, unlike the infant task, requires explicit knowledge, this should hold for both modalities. That is, since the same explicit choice procedure was used for both of our matching tasks, performance in both tasks should have been at chance. However, children in both age groups performed above chance on the visual-visual control task.

A second possible explanation is that the present findings together with those of the infant study provide evidence of a U-shaped developmental curve. One interpretation of U-shaped curves is that they arise from the conflict that occurs when a new representational system replaces a preceding one (Strauss, 1982). After the second system is established as the preferred approach, a higher level of performance is regained. If infants have a preverbal mechanism for representing numerosity, then one way U-shaped development could arise is that, as children acquire the conventional counting system, they attempt to count to solve problems even though they are not yet proficient counters. Of course, this explanation would not work if one assumes that a single infant mechanism applies to both auditory and visual sets, because then this U-shaped curve should be evident for both of the present tasks. However, one might posit that, rather than using one mechanism to determine numerosity in both modalities, infants use separate mechanisms that are modality specific. In this case, development could occur differentially for each mechanism. For example, U-shaped development might occur in the auditory modality and not the visual modality because the conventional counting system maps more directly onto the visual preverbal mechanism or because greater exposure to counting visual sets facilitates the mapping process. This particular explanation seems unlikely given our results indicating that counting proficiency is not required to perform the visual-visual task. However, this does not rule out the possibility that some other strategy emerges in this age period that affects the auditory mechanism differently than the visual mechanism.

A third possibility is that the abilities attributed to infants by Starkey et al. (1990) do not actually emerge until later in development. These authors have claimed 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 using one-to-one correspondence. This account implies that exact representations of number underlie infants' behavior. However, as Huttenlocher, Jordan, and Levine (1994) have argued, even if infants represent sets of discrete entities, these representations might be inexact for small set sizes, as are adults' representations for large set sizes if they do not count. Because the data reported in infant number studies are averaged over trials and over infants, differential looking times toward displays that vary in numerosity might occur on only some percentage of trials based on an approximating mechanism. In contrast to looking time measures, our task required exact detection of numerical correspondences. Thus, whereas the Starkey et al. (1990) procedure might have been tapping an approximate representation, the present tasks required an exact representation.

Finally, it is possible that the infant effect is not reliable. Starkey et al. (1990) based their interpretation on effects that were quite small. Furthermore, attempts to replicate these effects have yielded an inconsistent group of findings. When infants in the original study were shown pairs of visual displays with two and three objects and then presented with a set of either two or three drumbeats, they looked significantly longer toward the equivalent displays. However, subsequent replication attempts have reported that infants looked significantly longer toward the *nonequivalent* displays (Mix, Levine, & Huttenlocher, 1994; Moore, Benenson, Reznick, Peterson, & Kagan, 1987). The fact that existing studies of infants' detection of intermodal numerical correspondences report small effects in both directions of preference suggests the possibility that the results could have arisen by chance. That is, the published studies could be in the tails of a normal distribution of findings with many unpublished nonsignificant findings in the center. In this case, the present results would not be discrepant with the infant literature and would correspond with well-established findings of infant sensitivity to numerosity in visual sets (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981). Furthermore, preschooler's poor performance on the auditory-visual task in relation to the visual-visual task provides additional evidence for the bias reported in studies of counting system acquisition (Schaeffer et al., 1974; Shipley & Shepperson, 1990; Wynn, 1990).

The question of when the ability to detect intermodal numerical correspondences emerges has significant implications for theories of quantitative development. If infants have this ability, then it will be important to determine how it relates to subsequent developmental achievements, such as performing explicit choice tasks and acquiring the conventional counting system. If this competence is not present in infancy, then the question becomes how intermodal numerical matching ability develops in early childhood.

References

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

Gallistel, R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43–74.

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Essays on biology and cognition*, Hillsdale, NJ: Erlbaum.

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*. New York. Cambridge.

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

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

Mix, K. S., Levine, S. C., & Huttenlocher, J. *Infants' detection of auditory-visual numerical correspondences. Another look*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, May 1994.

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.

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology*, **6**, 357–379.

Shipley, E. F., & Shepperson, B. (1990). Countable entities: Developmental changes. *Cognition*, **34**, 109–136.

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

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.

Strauss, S. (1982). Introduction. In S. Strauss & R. Stavy (Eds.), *U-shaped behavioral growth*. New York: Academic Press.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, **36**, 155–193.