Child Development, September/October 2005, Volume 76, Number 5, Pages 978 – 988

Preschool Children's Mapping of Number Words to Nonsymbolic Numerosities

Jennifer S. Lipton and Elizabeth S. Spelke
Harvard University

Five-year-old children categorized as skilled versus unskilled counters were given verbal estimation and number word comprehension tasks with numerosities 20 – 120. Skilled counters showed a linear relation between number words and nonsymbolic numerosities. Unskilled counters showed the same linear relation for smaller numbers to which they could count, but not for larger number words. Further tasks indicated that unskilled counters failed even to correctly order large number words differing by a 2 : 1 ratio, whereas they performed well on this task with smaller numbers, and performed well on a nonsymbolic ordering task with the same numerosities. These findings provide evidence that large, approximate numerosity representations become linked to number words around the time that children learn to count to those words reliably.

Although infants represent and discriminate large numerosities (Lipton & Spelke, 2003; Xu & Spelke, 2000) and adults draw on large-number representations when they perform symbolic arithmetic (Dehaene, 1997; Gallistel & Gelman, 2000), it is not known whether preschool children use these nonverbal representations of number to master the symbolic number system. The present experiment investigates whether children link number words to nonsymbolic representations of numerosity before, during, or after they have learned to count to large numbers reliably.

Studies of human infants, adults, and animals provide evidence for a language-independent representation of approximate numerosity (Dehaene, 1997; Gallistel, 1990; Meck & Church, 1983). For example, human adults successfully compare sets of dots and sequences of tones when the items are too numerous and too briefly presented for verbal counting, with accuracy dependent upon the ratio of the two sets (Barth, Kanwisher, & Spelke, 2003; Van Oeffelen & Vos, 1982). Moreover, adults prevented from counting are able to press a key a specified number of times, with scalar variability in their responses proportional to numerosity (Whalen, Gallistel, & Gelman, 1999). These findings provide evidence for a language-independent system for representing numerosity with a Weber ratio limit as the signature of discrimination precision.

Turning to human infants, experiments provide evidence for discrimination of large numbers of dots and tones (Lipton & Spelke, 2003, 2004; Xu & Spelke, 2000). Number discrimination in infants also follows Weber's Law, such that the discriminability of two numbers depends upon their ratio (Lipton & Spelke, 2004; Xu, 2003). Whereas adults discriminate numerosities at about a 1.15 ratio (Van Oeffelen & Vos, 1982), 6-month-old infants discriminate numerosities at a 2.0 ratio (4 vs. 8 and 8. vs. 16 elements, but not 4 vs. 6 or 8 vs. 12), and 9-month-old infants discriminate numerosities at a 1.5 ratio (4 vs. 6 and 8. vs. 12 elements, but not 4 vs. 5 or 8 vs. 10) (Lipton & Spelke, 2003, 2004; Xu & Arriaga, in preparation). These findings suggest that numerical discrimination becomes more precise during infancy, prior to the onset of language, but remains less precise than that of adults.

In addition to an approximate, nonsymbolic representation of number, adults possess a uniquely human number representation that is linked to language. Behavioral and neuroimaging experiments have begun to shed light on the nature and interplay of these systems (see Dehaene, 1997, for general discussion). For example, experiments have revealed different patterns of brain activation when participants performed exact arithmetic (i.e. $4 + 5 = 9$ vs. 7) versus approximate arithmetic (i.e. $4 + 5 \approx 8$ vs. 3), with greater activation of secondary language areas for the former problems (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). Additionally, when

We thank Stanislas Dehaene, Susan Carey, and members of Harvard's Laboratory for Developmental Studies for advice and assistance. This research was supported by a Harvard University Fellowship to JSL and NSF grant REC-0087721 to ESS.

Correspondence concerning this article should be addressed to Elizabeth Spelke, Department of Psychology, Harvard University, 1120, William James Hall, 33 Kirkland Street, Cambridge, MA 02138. Electronic mail may be sent to spelke@wjh.harvard.edu.

 0009-3920/2005/7605-0003

bilingual adults were trained on both types of problems in one language and then tested on the same problems in their other language, there was a cost for switching languages for exact but not approximate arithmetic (Dehaene et al., 1999). These converging behavioral and neuroimaging results provide evidence for two systems: A language-independent, approximate system and a language-dependent system that stores learned facts. Skilled arithmetic performance depends on both these systems, for neurological patients show calculation deficits after damage to either system (Lemer, Dehaene, Spelke, & Cohen, 2003).

At around 3 years of age, children learn verbal counting and begin to gain an understanding of the exact, symbolic system of number representation (see Gelman & Gallistel, 1978). Surprisingly, learning the meaning of the number words is a protracted task for children. Children learn the meaning of the word ''one'' about 6 months before they learn the meaning of ''two'' and about 9 months before they learn the meaning of ''three'' (Wynn, 1992). Once toddlers have mastered the numbers up to about 4, however, they appear to grasp the logic of the system and exhibit an understanding that each word in their count list picks out a specific, unique cardinal value that is one higher than the value picked out by the prior word (Wynn, 1992). Over the next several years, children extend their count list to much larger numbers. We focus here on children's understanding of large-number words, both within and beyond their counting range.

Adults understand at least three aspects of the meanings of words for large numbers. First, each word picks out a specific, exact numerosity: The application of a particular number word to a set changes, therefore, if a single item is added to or removed from the set. Second, each word conveys an approximate numerical magnitude: When we are told that a person weighs 98 pounds, or that the temperature is 45°, or that a box contains 27 letters, we have a sense of the size of the person, the warmth of the air, and the contents of the box. Third, later words in the count list correspond to larger numbers than earlier words. Only recently has research begun to probe children's developing mastery of these aspects of number word meaning.

Recent research provides evidence that 5-year-old children, who understand number words within their counting range, also understand that number words beyond that range refer to specific exact numerosities. When children who cannot count beyond 50 are told that a jar contains ''eighty-six marbles,'' for example, they judge that the jar no longer contains ''eighty-six'' after a single marble is removed or added to the jar. Moreover, they judge that the jar has ''eighty-six'' marbles if one marble is removed and a different marble replaces it, but not if the original marble is then restored (Lipton & Spelke, in press). These findings provide evidence that preschool children have mastered one aspect of the logic of number word meanings and apply that logic to number words beyond their counting range.

Do such children also map number words onto nonsymbolic representations of number? In principle, children might begin to form these mappings before, during, or after learning to count to large numbers. In the first case, children might learn the mapping by forming direct associations between individual number words and nonsymbolic numerosity representations, prior to learning the count sequence. Before children learn that the word ''a hundred'' comes after ''ninety-nine,'' they might learn that the word ''hundred'' refers to approximately the number of marbles that would fill a vase, or approximately the temperature of a very hot day, and the like. These associative mappings then might guide the development of counting skills.

In the second case, children might acquire the mapping of number words to non-symbolic number representations at the same time as they learn to order those words appropriately in their count list. As children learn where to place each number word in counting, they might automatically and immediately map these words onto nonsymbolic numerosity representations. For example, children may learn to count to large numbers by enumerating visual arrays with many objects. As they produce each number word, they may associate it with the set of objects counted thus far. In the future, each number word would then call up a representation of the associated approximate numerosity, even without counting.

In the third case, children might come to map number words onto nonsymbolic numerosity representations only well after they have learned to count to those numbers. Children may initially learn the count list as simply a memorized list and a mechanical routine, without attaching any sense of numerical magnitudes to the words. At first blush, this alternative seems implausible, because number words are so laden with meaning for us as adults. As Giaquinto (personal communication, 2005) has noted, however, adults are apt to form this kind of representation when we learn new counting procedures. Adults who have learned binary counting, for example, can judge reliably that the number after 100,110 is 100,111. Many such adults have no sense, however, of the numerical magnitudes that these numbers convey: Is 100,110 °F a day for swimming or skiing? Is 100,110 years the age of a child, a young adult, or an octogenarian? The inability of many adults to answer such questions indicates that it is possible to master a symbolic counting routine without having any sense of the numerical magnitudes that individual symbols represent. Like such adults, children who have learned verbal counting initially may fail to map words like ''eighty-six'' to any nonsymbolic representation of numerosity.

Several findings bear on these possibilities. Huntley-Fenner (2001) asked 5 – to 7-year-old children to give verbal numerical estimates of visual arrays containing 5 – 11 items. Children's performance was similar to that of adults, showing both the Weber signature and scalar variability (Huntley-Fenner, 2001). These findings suggest preschool children have mapped number words up to 11 onto their approximate number representations. In another study, 5-year-old children and adults were asked to judge whether a target numerosity, presented either as an Arabic numeral or as an array of dots, was greater or less than 5 (Temple & Posner, 1998). Both adults and children demonstrated a distance effect, i.e., better performance for numbers far from 5 than numbers close to 5. This effect has been interpreted as stemming from a mapping between number symbols and an underlying representation of quantity (Dehaene, 1997). Event-related potential recordings during this task indicated, moreover, that similar brain regions were activated in children and adults when making these numerical judgments, suggesting a common mechanism of numerical estimation in adults and preschool children (Temple & Posner, 1998).

Finally, in a set of recent studies, children were asked to indicate where a given number, presented both as a number word and as an Arabic symbol, fell on a number line (Siegler & Booth, 2004; Siegler & Opfer, 2003). Children's performance on this task demonstrated a monotonic relationship between number words and positions on the line at both 5 and at 8 years of age (Siegler & Opfer, 2003). There was a qualitative change in this function, moreover, with younger children showing the logarithmic relation of number symbols to quantity predicted by Dehaene (1997) and older children showing the linear relation required for accurate arithmetic (Siegler & Booth, 2004). Children's performance correlated strongly with math achievement (Siegler & Booth, 2004).

Although these findings provide evidence for similarities between children's and adults' representations of number, they have several limitations. Importantly, none of the studies assessed children's counting abilities, and so they cannot reveal when children's mapping of number words onto nonsymbolic representations occurs relative to the development of counting skills. Furthermore, the numbers tested in Huntley-Fenner's (2001) and Temple and Posner's (1998) studies were small (5 – 11 and 1 – 9, respectively) and likely to be well known to the children for over a year. Although Siegler and Opfer (2003) asked children about larger numbers, they did not test how well children could count to these numbers, and their task required children to map a numerical quantity onto a line. Therefore, it remains unclear when children learn to map words for large numbers onto nonsymbolic sets.

The current study addresses this question. We asked 5-year-old children and adults to assign number words to large sets of objects without counting in three tasks: (1) an estimation task in which subjects are given an array of elements and asked to produce a numerical estimate of its cardinal value, (2) a comprehension task in which subjects are given a number word and are asked to choose the corresponding array of elements, when that array is paired with an array that is half or twice as numerous, and (3) an ordering task in which subjects are shown two sets that differ in a 2 : 1 ratio, are told the cardinal value of one set, and are asked to estimate the value of the other set and are scored for choice of a larger or smaller number word. Before performing these tasks, children's skill at counting numbers up to 100 was assessed, and children were classified as skilled or unskilled counters based on their counting ability within this range. If the mapping of number words onto nonsymbolic representations occurs prior to learning to count to those number words, children of all counting abilities may succeed on these tasks. If the mapping occurs as soon as children learn to count to them, then skilled counters should succeed on the tasks, and unskilled counters should succeed only with the number words to which they count reliably. Finally, if the mapping does not occur until well after children master large-number counting, then children might fail our tasks, regardless of their counting ability.

Participants

Participants were 27 preschool children (mean age, 5 years 6 months; range 5 – 0 to 6 – 2) and 24 adults (mean 24 years 9 months; range 18 to 49 years). Three additional children were excluded for refusing to complete at least half of the experiment. Children were recruited from the participant lists of the MIT Infant Cognition Laboratory and Harvard Laboratory for Developmental Studies; adults were students or staff in the Harvard/MIT community. Children were recruited from a database compiled from birth records and mass mailings. The demographic characteristics of our participant pool were 60% White, 23% unknown, 11% more than one race, 4% Asian, 2% Black or African American, and <1% Native American or Alaska Native.

Counting Assessment

For the child participants only, the experiment began with an assessment of children's skill at counting to 100.

Method

At the start of the session, children were asked for the highest number to which they could count. If a child began to count spontaneously, his or her counting sequence was recorded. Then the experimenter began counting and children were asked to continue counting after the experimenter stopped. For example, the experimenter said, "55, 56, 57" and the child was asked to continue counting from 57. The experimenter told him or her to stop counting after the child had reached the decade change. The three sequences the experimenter asked every child were "55, 56, 57", "78, 79", and "95, 96." These items were chosen to both examine whether the child was able to add one to the units place, and to elicit whether the child was able to make the decade change (e.g., from 79 to 80) correctly. If a child stated that the highest number he or she could count to was a number that was much smaller than 50, the experimenter began with smaller numbers such as "23, 24, 25." If a child failed at any of the decade changes, the experimenter also asked children to count these smaller numbers.

Each participant was classified as either a "skilled counter" or an "unskilled counter" based on his or her performance on this task. If a child responded with perfect counting, the child was classified as a skilled counter. If a child made mistakes on any of the decade changes tested (e.g., said "58, 59, 30" or "50–10"), he or she was classified as an unskilled counter. If a child self-corrected, he or she was given credit for the correction. If a child appeared to be "close" to a correct answer, more counting was elicited to determine his or her classification. On the rare occasions when classification was unclear, the counting test was repeated at the end of the experiment. Children who said that they did not know how to continue counting, and refused to guess, were classified as unskilled counters.

Results

In total, 15 children were classified as skilled counters (mean age 5–6, range 5–0 to 6–2) and 12 as unskilled counters (mean age 5–4, range 5–0 to 5–11). All children in the latter category made multiple mistakes or omissions on the counting task, including at least two mistakes on the three tested decade changes (at 60, 80, and 100). These findings suggest that 5-year-old children differ considerably in their abilities to count to 100. The performance of the skilled and unskilled counters is assessed separately in all the tasks that follow.

Estimation

The second task assessed the ability of adults and children to verbally estimate the cardinal value of a large set without counting. Past research with adults provides evidence that numerical estimates are linearly related to the number of elements presented (e.g., Cordes, Gelman, Gallistel, & Whalen, 2001; Whalen et al., 1999), that the variability in numerical estimates is proportional to numerical magnitude (Cordes et al., 2001; Huntley-Fenner, 2001; Platt & Johnson, 1971; Whalen et al., 1999), and that round numbers such as multiples of 10 or 25 are offered disproportionately as estimates (Dehaene, Dupoux, & Mehler, 1990; Dehaene & Mehler, 1992). Because children will not typically tolerate enough trials to calculate variability estimates, we tested for the first and third properties of estimation performance within each of the 3 groups of participants (adults, skilled counters, and unskilled counters).

Method

Participants were first presented with two cards, one with a single circle and one with 300 circles, and were told: "Here is a card with one circle on it. Here is a card with 300 circles: It has too many circles for you to count, but there are 300 circles on the card. I am going to show you cards with more than this [pointing to card with one] and less than this [pointing to card with 300]." Then children were asked to estimate how many items were on each of a set of cards, presented too briefly for counting. The estimation task consisted of 12 trials with large numerosities 20–120 (experimental) and three trials with numerosities under 10 (control). Children received additional control trials with small num-

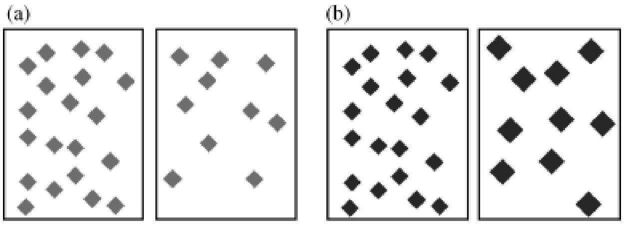

Figure 1. Sample displays used in the four tasks. For the estimation task, a single display was presented on each trial; for the other tasks, two displays of numbers differing by a 2 : 1 ratio were presented, matched in either (a) element size or (b) summed area.

on this and subsequent tasks. Therefore, the performance of a subset of children was recoded by coders who were unaware of the counting performance of each participant and also unaware of the correct answers on this and subsequent tasks. Of the 32 participants, 15 were recoded (9 of the 17 skilled counters and 6 of the 15 unskilled counters), and the correlations between the initial coding and the naïve coding were computed. Inter-rater reliabilities of the verbal estimates for the 15 children recoded, and for the 9 skilled counters and 6 unskilled counters separately, were all high, with R^2s = .992, .9995, and .982, respectively, all ps < .001.

erosities, whenever the experimenter suspected that they were beginning to respond at random. On each trial, participants were shown a card displaying a total of 4, 6, 7, 20, 40, 60, 80, 100, or 120 pink diamonds. As a partial control for the continuous variables of element size and total filled area of the display, there were two sets of displays: A set in which element size was constant and a set in which total summed area was constant across the different numerosities. In the former set, the more numerous arrays had greater total summed area; in the latter set, they had smaller element sizes. Figure 1 presents examples of each type of display.

Because the experimenters were aware of the counting ability of each child and the correct answers for some trials, it is possible that their knowledge could influence their coding of children's responses

Trials were presented in a quasi-random order such that trials with the same numerosity were never consecutive. The three control questions were presented about every 3–5 trials to ensure that participants were on-task. If a participant provided a numerical answer that was larger than 300, the cards with 1 and 300 circles were presented again and he or she was reminded that all cards in the game had less than 300 on them. This never occurred for adults, occurred on one trial for 1 skilled counter, and occurred on eight trials for 4 unskilled counters.

Results

Figure 2a presents the means and standard errors of adults' verbal estimates for each numerosity. Adults were highly accurate at estimating numero-

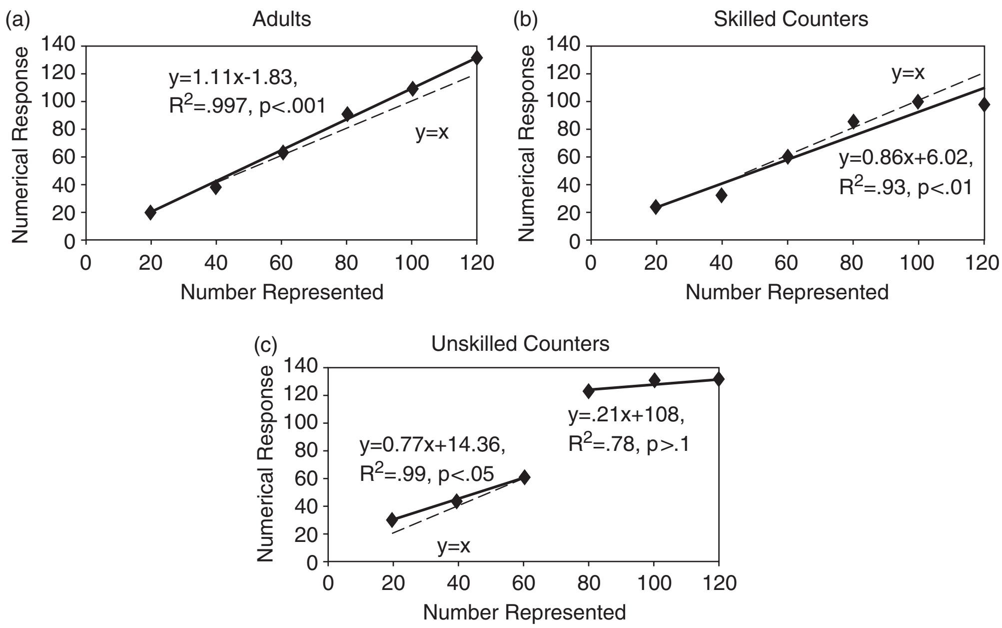

Figure 2. Estimation performance (diamonds), best-fitting linear regressions (solid lines) and ideal performance (dotted lines) for (a) adults, (b) children categorized as skilled counters, and (c) children categorized as unskilled counters. In (c), regressions are calculated separately for numbers within versus outside children's counting range.

sity, and their estimates increased linearly with increasing numerosity ($y = 1.11x - 1.83$, $R^2 = .997$, $p < .001$). Adults' estimates did not differ for the area constant versus element size constant sets, $F(1, 23) < 1$.

Figure 2b presents the same data for the 5-year-old skilled counters. These children's estimates also were linearly related to the presented numerosity ($y = 0.86x + 6.02$, $R^2 = .93$, $p < .01$) and did not differ for the area- versus element-constant sets, $F(1, 14) = 2.07$, $p > .10$.

Figure 2c presents the same estimation data for the 5-year-old unskilled counters. Although children tended to provide larger number words for larger sets, the slope of the linear regression analysis was not significantly different from zero ($y = 1.19x + 3.68$, $R^2 = .90$, $p > .10$). Estimates again did not differ for the area constant and element size constant sets, $F(1, 11) < 1$.

Because all of the children categorized as unskilled counters succeeded at counting with low numbers, we analyzed their estimation data separately for trials testing the numbers 20–60 and 80–120, as an initial test for differences between estimates within versus outside children's counting range. Figure 2c presents the numerical estimates for the two sets of numerosities. For numerosities 20–60, the estimates of unskilled counters increased linearly with increasing numerosity ($y = 0.77x + 14.36$, $R^2 = .99$, $p < .05$); for the numbers 80–120, the estimates did not vary as a function of the presented numerosity ($y = 0.21x + 108$, $R^2 = .78$, $p > .10$).

Finally, we tested whether participants spontaneously produced round number words for large numerosities. First, we separated responses on the 12 experimental trials into three categories: Numbers less than 10, multiples of 10, and all other responses. Figure 3 presents the percentage of numerical responses in each of these categories for the 3 groups of participants. If participants chose numbers randomly, their answers should be multiples of 10 only 10% of the time. As Figure 3 indicates, responses that were multiples of 10 were produced well above chance for each group of participants: $t(23) = 13.97$ for adults, $t(14) = 6.29$ for skilled counters, and $t(11) = 5.97$ for unskilled counters, all ps $< .001$, and there was no difference among the groups, $F(2, 48) = 3.01$, $p > .05$.

Discussion

Both adults and 5-year-old skilled counters were able to estimate large numbers, and their estimates followed the signature linear function found in past research (e.g., Cordes et al., 2001). Five-year-old children who were unskilled counters also gave estimates showing a linear function for numerosities within their count range. Outside their counting range, however, the estimates of unskilled counters were unrelated to numerosity, although they were higher overall than estimates for numbers within their counting range. For all 3 groups, there was no effect of continuous quantities on numerical estimates. Finally, all 3 groups tended to give round numbers as responses, as in past research with adults (Dehaene et al., 1990).

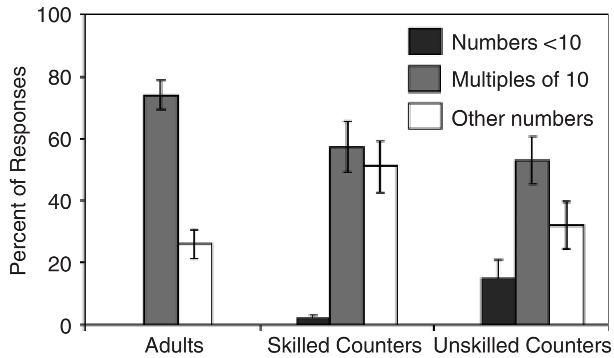

Figure 3. Patterns of usage of different types of number words in the estimation task.

Because the children in this experiment were 5 years old and few had begun formal schooling, these findings suggest that children quickly learn to map the words in their count sequence onto nonsymbolic numerosities. The next task was undertaken to test that suggestion further, by investigating whether adults, skilled counters, and unskilled counters can select the nonsymbolic numerosity that better corresponds to a presented number word.

Number Word Comprehension

The comprehension task assessed whether children would match a presented, large number word to an array with the appropriate number of elements, over an array with twice or half as many elements.

Method

The comprehension task immediately followed the estimation task. On each trial, participants were shown two cards similar to those used in the estimation task, except that the cards presented rectangles instead of diamonds. One card—the target—displayed 20, 40, 60, 80, 100, or 120 rectangles. The other card—the distractor—presented half as many elements on half the trials and twice as many elements on the remaining trials. On one trial with each pair of numbers, the two cards presented rectangles of the same size; on another trial with each pair, the two cards presented the same amount of filled area (and so the elements were smaller in the card presenting the larger number). Cards were presented so that the target appeared on the left on half the trials. After a participant had looked at both cards, the experimenter asked him or her to point to the card with "N" shapes, where N corresponded to the target numerosity. Each participant received 24 test trials and 8 control trials, in which a large-number array was paired with a small-number array (e.g., 7 vs. 50). Control trials were interspersed with the test questions to maintain motivation and to check that children were on-task. Participants were given informative feedback after each response.

Children's and adults' responses were coded live by the experimenter, and a subset of 15 children's responses were recoded by observers who were blind both to the correct answer and to the child's counting ability. Reliabilities between the live and naïve coders were all high, with R^2s = .995, .999, and .987, all ps < .001, for the 15 participants, the 9 skilled counters, and the 6 unskilled counters, respectively.

Results

Figure 4 presents the data for the 3 groups of participants. Overall, adults scored significantly above chance ($t(23) = 33.89$, $p < .001$) with no difference between performance on the trials with constant element size versus area ($t(23) = 0$), or between performance when the correct answer was the bigger or smaller set ($t(23) = 1.64$, $p > .10$).

The 5-year-old skilled counters also scored well above chance ($t(14) = 5.81$, $p < .001$), with marginally better performance when summed area was constant ($t(14) = 2.13$, $p = .052$), and with better performance when the correct answer was the bigger set, $t(14) = 3.17$, $p < .01$.

Because only 8 of 12 unskilled counters completed more than half of this task, the analysis included only their data. Overall, these children performed at chance, $t(7) = 1.83$, $p > .10$, with no difference between performance on constant element size and constant area sets, $t(7) = 1.87$, $p = .10$, but with better performance when the correct answer was the larger set, $t(7) = 4.25$, $p < .01$. In general, these children tended to select the set with the greater numerosity regardless of the number word presented. Analyses comparing these children's performance with smaller versus larger number words yielded no effect of this variable: Children performed at chance both

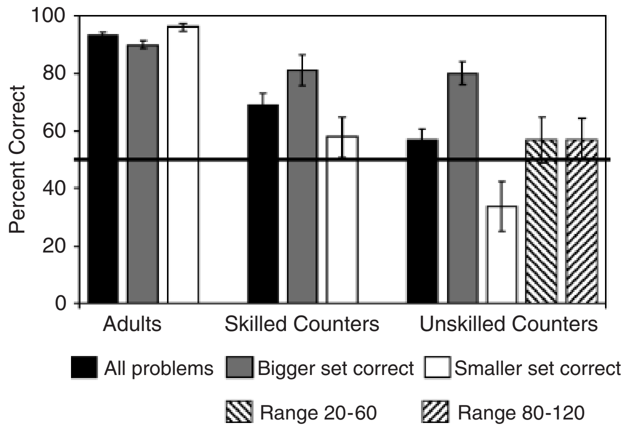

Figure 4. Adults' and children's performance on the comparison task when the correct answer was the more or less numerous array and when the target was within or outside the unskilled counters' counting range.

with numbers in the range of 20–60 (mean = 57%, $t(7) = 1.22$, $p > .05$), and with numbers in the range of 80–120 (mean = 57%, $t(7) = 2.30$, $p > .05$).

The performance of the 3 groups of participants was compared using a 3 Group (Adults, Skilled Counters, Unskilled Counters) × 2 Pair Type (element-controlled or area-controlled) × 2 Size (Larger vs. Smaller array the correct answer) ANOVA. This analysis revealed a reliable main effect of Group, $F(2, 44) = 52.5$, $p < .001$, with adults performing the best and unskilled counters performing the worst. A post hoc Tukey test showed that adults performed better than skilled and unskilled counters, and skilled counters performed better than unskilled counters, with all ps < .01. Furthermore, performance was significantly better when the summed area was controlled than when the element size was controlled, $F(1, 44) = 7.53$, $p < .01$. There was also a significant effect of Size (i.e. better performance when the correct answer was the bigger set), $F(1, 44) = 31.42$, $p < .001$, that interacted with Group, $F(2, 44) = 16.73$, $p < .001$. Adults performed better than both skilled and unskilled counters when the smaller set was correct, and skilled counters performed better than unskilled counters when the smaller set was correct (all ps < .05, Tukey post hoc tests). Finally, there was an interaction between Pair Type and Size $F(1, 44) = 7.01$, $p < .05$, with better performance when the smaller set was correct on the summed area constant trials.

Discussion

The findings of the comprehension task provide further evidence that children map number words onto representations of nonsymbolic, large numerosities when they become proficient at counting. Both adults and skilled counters were able to match presented number words to arrays of appropriate numerosity, in preference to arrays that were twice or half as numerous. In contrast, unskilled counters failed this matching task. The findings of the number word comprehension task therefore converge with those of the estimation task in suggesting a change in children's mapping of number words to nonsymbolic numerosities, at about the time that children master counting to large numbers.

Number Word Ordering

The next task investigated whether children understand that higher number words pick out larger numerosities than lower number words. Children were shown two arrays whose numerosities differed by a 2.0 ratio, they were given the verbal cardinal value of one array, and then they were asked to estimate the number of elements in the other array. When children were told (e.g.) that the more numerous array contained ''eighty'' elements, would they choose a smaller number word for the less numerous array?

Method

Immediately following each trial of the above comprehension task, children were given appropriate feedback: They were told which of the two arrays contained the stated number of elements. Then children were given a second estimation task: They were asked how many elements appeared on the other card (the one not labeled). Because the distractor card on each trial contained either half or twice as many elements as the target card, children's estimates for the unlabeled card can be analyzed in two ways. First, we may ask whether children produced a number word in the correct direction: If the unlabeled card presented a more numerous array, did children produce a word for a higher numerosity than the word that the experimenter presented and applied to the labeled card? Second, we may ask how closely both children and adults approximated the correct numerical label: When the unlabeled card contained twice as many elements, did participants select a number word for a value twice as large as the presented number word?

As in the first estimation task, children's responses were recorded by the experimenter, and the responses of a subset of children were recoded by observers who were blind to the child's counting status and to the correct answer on each trial. Reliabilities between the original and the blind coding were high, for all 15 participants considered together, and separately for the skilled counters and unskilled counters, with R^2s = .958, .961, and .951, respectively (all ps $<$ .001).

Results

Figure 5 shows the percentage of trials on which participants provided an answer in the correct direction for each of the labeled numerosities. As expected, adults always provided answers in the correct direction, producing a number word that was larger than the identified numerosity when the target numerosity was larger and a smaller number word when the target numerosity was smaller. For the adults, there was a highly significant linear relationship between numerical estimates and number presented ($y = 0.94 + 2.80$, $R^2 = .999$, $p < .001$; Figure 6).

The children who were skilled counters also produced answers in the correct direction on almost all trials, $t(14) = 22.6$, $p < .001$ and showed a linear relationship between numerical estimates and number presented ($y = 0.54 + 18.76$, $R^2 = .975$, $p < .001$). As on the first estimation task, however, the slope of this relationship showed a tendency to underestimate numerosity.

Many of the unskilled counters refused to produce number words on this task, but those who did respond produced an estimate in the correct direction significantly above chance $t(7) = 3.55$, $p < .01$. A further analysis considered their responses separately for the numerosities that were within their count range (20–60) versus outside that range (80–120). These children were able to provide answers in the correct direction for the smaller numerosities ($t(7) = 8.64$, $p < .001$), but not for the larger num-

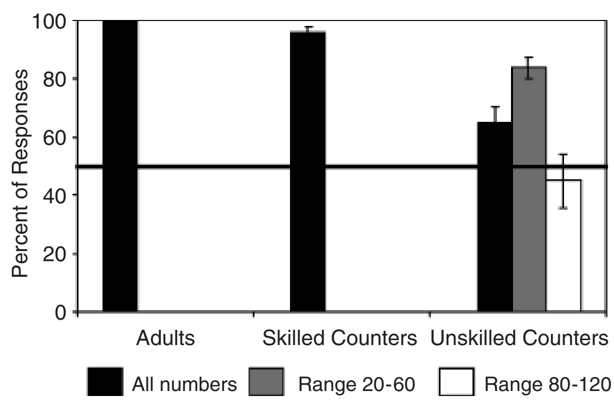

Figure 5. Percent of responses in the correct direction on the number word ordering task for adults, skilled counters, and unskilled counters tested with words within versus outside their counting range.

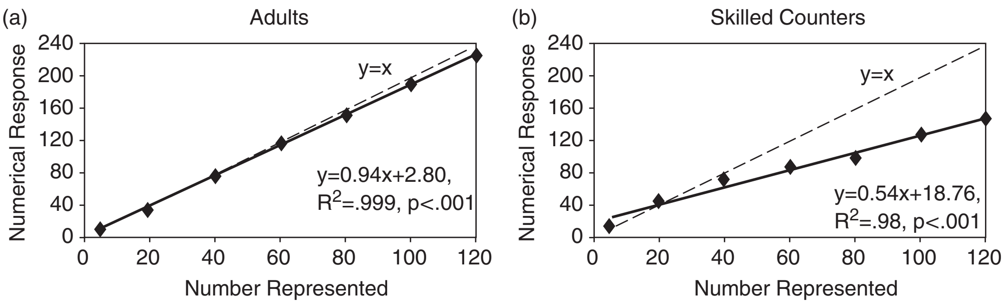

Figure 6. Performance on the number word-ordering task (diamonds), best-fitting linear regressions (solid lines), and ideal performance (dotted lines) for (a) adults and (b) children categorized as skilled counters.

erosities ($t(7) < 1$); performance on these two ranges differed reliably, $t(7) = 2.82$, $p < .05$. There was insufficient data to perform linear regressions on the unskilled counters' estimations.

Discussion

When given information about one set, both adults and children who were skilled counters successfully used this information to provide a numerical estimate in the correct direction for the unidentified set. Unskilled counters also gave numerical estimates in the correct direction for numbers within their counting range. In contrast, unskilled counters randomly gave larger versus smaller number word estimates for numbers outside their counting range. These findings provide evidence that soon after children learn to place number words in the count list, they map these words onto nonsymbolic quantity representations that capture the ordinal relationships between the number words.

Nonsymbolic Numerosity Discrimination

When children or adults appropriately pair a number word with a set of nonsymbolic numerosities, their performance ensures that they perceive the nonsymbolic numerosities, understand the words, and map these representations to one another. When a participant fails appropriately to pair number words with nonsymbolic numerosities, however, that failure could stem either from failure to discriminate the nonsymbolic numerosities or from failure to construct the appropriate mapping. The final task was undertaken to distinguish the first and last sources of failure.

Method

Immediately following the number word ordering task, participants were shown pairs of cards with green diamonds side by side, one with twice as many diamonds as the other, and they were asked to point to the card with "more diamonds." Except for the change in wording (particularly the absence of number words) and in element color and shape (undertaken to maintain children's interest), the structure of this experiment was the same as that of the comprehension task. Participants were given 10 discrimination trials (8 test and 2 control) with a subset of the pairs of numerosities used in the previous tasks. Half the trials presented displays with equal element sizes (10 vs. 20, 40 vs. 80, 50 vs. 100, and 120 vs. 240 and control trial 6 vs. 40) and half presented displays with equal summed area (10 vs. 20, 20 vs. 40, 30 vs. 60, and 80 vs. 160 and control trial 7 vs. 60). The side of the correct answer was counterbalanced.

Children's choices were recorded by the experimenter, and the responses of a subset of children were recoded by observers who were blind to the child's counting status and to the correct answer on each trial. Because of the lack of variance, reliabilities for the groups are not meaningful, but the percentage agreement between the coders was 99%.

Results and Discussion

Adults and 5-year-old skilled counters performed perfectly on this task (both *SD*s = 0), and unskilled counters performed well above chance (98% accuracy, $t(11) = 19$, $p < .001$). The high performance shown by all participants indicates clearly that errors in the symbolic tasks stemmed from either lack of knowledge of the number words or errors in the mapping from number words to nonsymbolic numerosities, not in the discrimination or comparison of the nonsymbolic numerosities.

General Discussion

The present experiment investigated whether preschool children begin to map words for large numbers onto nonsymbolic number representations before or after the acquisition of skilled verbal counting. Three tasks testing numerical estimation, number word comprehension, and number word ordering provide two kinds of evidence that children construct this mapping during or soon after learning the count sequence. First, children who cannot count reliably beyond 60 fail the estimation, comprehension, and ordering tasks for number words larger than 60, whereas adults and children who can count to 100 pass these tasks for all number words. Second, children who fail to count reliably to 100 succeed at the estimation and ordering tasks when tested with number words within their counting range but fail for larger number words.

It is possible that the superior performance of skilled counters, relative to unskilled counters, depends in part on extraneous variables that distinguish these children, such as intelligence, motivation, or maturity. Such variables cannot, however, account for the differential performance of unskilled counters when they are tested on number words within versus outside their counting range. Because the same children performed differently with numbers to which they could or could not count, the association between counting skill and success on our three tasks cannot be attributed to extraneous distinctions among our participants. Instead, children's success at mapping number words to nonsymbolic numerosities depends specifically on their success at counting to those words.

Our data suggest, therefore, that children form the mapping between symbolic and nonsymbolic representations of number at about the time they master the count sequence. These findings may help to explain why we found no children who had partly mastered the counting sequence from 60 to 100. Although children learn very slowly the meanings of the first four number words (Wynn, 1992), and most children do not learn the meaning of the counting routine until the end of the 4th year, over half of the 5-year-old children that we tested knew the count sequence up to 100. Among the remaining 5-year-old children, we found none who could count to 80 but not to 100. Although such children may exist, their rarity suggests that children master these words very rapidly. If children map each new word in their count list to an appropriate nonsymbolic numerosity representation, then this mapping may provide children with a stable system of representation to anchor new number word meanings, supporting their rapid learning. Nevertheless, the present studies do not reveal whether children's mapping of new counting words to numerosity representations occurs immediately upon mastery of counting, or occurs after delays too brief to be detected with the present methods. Training studies, in which children are taught to extend their count list, are needed to address this question.

When children begin school, many have difficulty learning elementary arithmetic that is taught by rote, with a curriculum that begins with the smallest numbers and insists on exact, accurate calculations. In contrast, recent research shows that preschool children not only discriminate numerosities in the range of 20–120, as they did in the last task presented here, but they also can add such numerosities and compare their sums to contrasting numerosities before learning symbolic arithmetic (Barth et al., in press; Barth, LaMont, Lipton, & Spelke, in review). The present evidence that children spontaneously map number words to nonsymbolic approximate numerosities therefore invites a new look at the elementary arithmetic curriculum. Mathematics educators may be able to harness both preschool children's understanding of nonsymbolic arithmetic and their spontaneous mapping between nonsymbolic and symbolic number systems, to enhance elementary-school children's learning of arithmetic.

References

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representation in adults. *Cognition*, *86*, 201–221.

Barth, H., LaMont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. S. (in press). Nonsymbolic arithmetic in adults and preschool children. *Cognition*.

Barth, H., LaMont, K., Lipton, J., & Spelke, E. S. (2005). Numerical computation and abstraction in preschool children. Manuscript submitted for review.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, *8*, 698–707.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 626–641.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29.

Dehaene, S., Spelke, E. S., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970–974.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Science*, *4*, 59–65.

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

Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': Numerical estimation by 5–7-year-olds. *Cognition*, *78*, B27–B40.

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*, 1942–1958.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science*, *14*, 396–401.

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, *5*, 271–290.

Lipton, J. S., & Spelke, E. S. (in press). Preschool children master the logic of number words. *Cognition*.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 320–334.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogenous behavior chain: Effects of error contingencies. *Learning and Motivation*, *2*, 386–414.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237–243.

Temple, E., & Posner, M. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Science*, *95*, 7836–7841.

Van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*, 163–170.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*, 130–137.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220–251.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*, B15–B25.

Xu, F., & Arriaga, R. (2005). Large number discrimination in 10-month-old infants. Unpublished manuscript.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.