
Volume 24, issue 3, July–September 2009 ISSN 0885-2014
COGNITIVE DEVELOPMENT
DEANNA KUHN
EDITOR
The Official Journal of The Jean Piaget Society for the Study of Knowledge and Development

Cognitive Development 24 (2009) 248–264

Contents lists available at ScienceDirect

Cognitive Development

Children's mappings of large number words to numerosities

Hilary Barth*, Ariel Starr, Jessica Sullivan

Department of Psychology, Wesleyan University, United States

ARTICLE INFO

Keywords:
Number
Numerical cognition
Counting

ABSTRACT

Previous studies have suggested that children's learning of the relation between number words and approximate numerosities depends on their verbal counting ability, and that children exhibit no knowledge of mappings between number words and approximate numerical magnitudes for number words outside their productive verbal counting range. In the present study we used a numerical estimation task to explore children's knowledge of these mappings. We classified children as Level 1 counters (those unable to produce a verbal count list up to 35), Level 2 counters (those who were able to count to 35 but not 60) and Level 3 counters (those who counted to 60 or above) and asked children to estimate the number of items on a card. Although the accuracy of children's estimates depended on counting ability, children at all counting skill levels produced estimates that increased linearly in proportion to the target number, for numerosities both within and beyond their counting range. This result was obtained at the group level (Experiment 1) and at the level of individual children (Experiment 2). These findings provide evidence that even the least skilled counters do exhibit some knowledge of the form of the mapping between large number words and approximate numerosities.

Converging evidence from hundreds of studies with humans and nonhuman animals demonstrates that the ability to represent and manipulate numerical quantity is not dependent on language or formal training (Barth, La Mont, Lipton, & Spelke, 2005; Dehaene, 1997; Gallistel & Gelman, 1992; Gilmore, McCarthy, & Spelke, 2007; Wynn, 1998). Humans' number sense is grounded in preverbal representations of number that we share with many nonhuman animal species (Dehaene, 1997; Gallistel & Gelman, 1992), and these nonverbal representations give us a variety of approximate numerical abili-

* Corresponding author at: Department of Psychology, Wesleyan University, 207 High St, Middletown, CT 06459, United States.
E-mail address: hbarth@wesleyan.edu (H. Barth).

0885-2014/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2009.04.001

ties. Both rats trained to press a lever a fixed number of times and humans told to press a key a fixed number of times as quickly as possible and without counting, for example, will produce sequences of presses that are normally distributed around the target number (Platt & Johnson, 1971; Whalen, Gallistel, & Gelman, 1999). Human infants attend to the numerical properties of stimuli spontaneously as well. Six-month-old infants in carefully controlled habituation studies are capable of discriminating between 4 versus 8 and 8 versus 16 dots or tones on the basis of numerosity, although they are unable to discriminate 8 versus 12 (Lipton & Spelke, 2003, 2004; Xu, 2003; Xu & Spelke, 2000). Infants' numerical discrimination patterns demonstrate the fundamentally approximate nature of preverbal number representations: at 6 months, infants successfully discriminate sets that differ by a numerical ratio of 2.0 and fail to discriminate sets that differ by a ratio of 1.5 (for a review, see Cordes & Brannon, 2008).

Preverbal infants and nonhuman animals appear to process numerical quantities through the use of mental representations in the form of continuous magnitudes (often called "mental magnitudes" or "analog magnitudes"). These mental magnitudes represent discrete quantities in the external world much like the continuous length of a column in a graph could represent the discrete number of children in a school district. Mental magnitudes underlie adults' numerical processing as well. Even adult humans with sophisticated numerical abilities draw on these preverbal approximate representations of number when estimating the numerosity of a set (Huntley-Fenner, 2001; Lipton & Spelke, 2005; Siegler & Opfer, 2003; Whalen et al., 1999) and when performing approximate calculations (Barth et al., 2006; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Spelke & Tsivkin, 2001).

Humans are not limited to representing number approximately. Language endows humans with the ability to represent the natural numbers through verbal count lists (Carey, 2001; Condry & Spelke, 2008; Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004; Hauser & Spelke, 2004; LeCorre & Carey, 2007; Sarnecka & Carey, 2008; but see Gallistel, 2007; Gallistel & Gelman, 1992, 2000). Natural number representations do not appear to emerge spontaneously, however. Both very young children and adults from populations without verbal count lists are unable to represent the natural numbers (Carey, 2001, 2004; Gordon, 2004; Hauser & Spelke, 2004; Pica, Lemer, Izard, & Dehaene, 2004; Wynn, 1992). The Amazonian Pirahã and Mundurukú tribes lack explicit number words for numbers above two and five, respectively, and do not have successive counting lists. While they are still able to manipulate approximate numerosities, they are unable to represent exact numerosities for sets above four or five (Gordon, 2004; Pica et al., 2004; but see Frank, Everett, Fedorenko, & Gibson, 2008). Thus, the verbal count list provides humans with an external means for representing large, exact numbers (Carey, 2001, 2004; Hauser & Spelke, 2004).

Adults in cultures with verbal count lists possess knowledge of diverse aspects of the number words and the verbal counting system. They have mastered the words in their particular list, or at least those within the range of typical experience. They understand the principles that govern the count list's representation of numerosity, and they can use those principles to assess the numerosity of a set of items. They know that every number word represents a specific, exact numerosity (even if it is a number word they have never encountered before), and they know that number words occurring later in the sequence refer to larger sets. Further, they have formed a mapping between the words in the verbal count list and the approximate numerosity representations of the mental magnitude system. Numerous behavioral and neuroimaging studies demonstrate that verbal number words and Arabic numerals are mapped to their corresponding semantic magnitude representations in adults (Dehaene et al., 1999; Huntley-Fenner, 2001; Lipton & Spelke, 2005; Moyer & Landauer, 1967; Whalen et al., 1999). How do we come to form this mapping between the uniquely human learned verbal count list and the phylogenetically widespread approximate magnitude system?

Opinions differ as to the role of the approximate magnitude system in children's initial acquisition of number word meanings. Some researchers have suggested that children initially attach meanings to number words by forming bidirectional mappings between number words and mental magnitudes (Gallistel, 2007; Gallistel & Gelman, 1992, 2000). Others have suggested that children acquire both number word meanings and basic counting principles through such a mapping (Wynn, 1995), or that they do so through the use of the mental magnitude system in addition to other representational capacities (Hauser & Spelke, 2004; Spelke, 2003; Spelke & Tsivkin, 2001). However, recent experiments provide evidence that the initial acquisition of verbal counting may not involve the mental magnitude system at all (LeCorre & Carey, 2007). These studies suggest that children begin to map the mental magnitude system's approximate numerical representations onto words in their verbal count list only about 6 months after they have constructed an understanding of counting principles, at the age of roughly four and a half (LeCorre & Carey, 2007). By age five, children have mapped analog magnitudes onto the number words up to at least 11 (Huntley-Fenner, 2001).

Because we are concerned with the acquisition of mappings to analog magnitudes, we focus here on children who have already come to understand the logic of counting in order to examine children's knowledge of mappings between number words and approximate numerical magnitudes, and to explore how that knowledge depends on children's mastery of verbal counting. Previous studies have shown that children do understand some properties of large number words that fall outside their productive counting range. Lipton and Spelke (2006) showed that 5-year-olds know that large number words outside their counting range apply to specific, exact numerosities. These children understand that a number-relevant transformation of a large set of objects (i.e. addition or subtraction) changes the number word that should be mapped onto the set and that a number-irrelevant transformation does not have this effect. This knowledge applies both within and outside the child's productive counting range, so it cannot depend upon prior mastery of the verbal count list.

In a different series of studies, however, Lipton and Spelke (2005) argued that mastery of the verbal counting list does precede children's understanding of another property of large number words: their mappings to approximate magnitudes. Five-year-old children who could reliably produce correct decade changes up to 100 (such as "78, 79, 80") could also produce increasingly larger verbal estimates for larger sets of items, and the estimates they produced were, on average, accurate (that is, they appeared to have mapped the number words in their counting range to approximate magnitudes). Children who had not mastered the counting sequence between 60 and 100, in contrast, failed to produce larger numerical estimates for larger sets outside their counting range. Thus, 5-year-old children appeared to produce estimates related to the presented numerosities only for sets that they were capable of counting (if given adequate time; during the study, children could not count the sets). No children appeared to have a partial knowledge of the verbal count list outside the range they had mastered. Instead, children showed evidence of a relatively accurate mapping of number words to analog magnitudes, but only after mastery of the verbal count list in the tested range. These findings led the authors to suggest that mastery of the verbal count list must precede mappings between number words and approximate numerical magnitudes, and that children might rapidly map each new learned word to an approximate numerosity (Lipton & Spelke, 2005).

Here we use a numerical estimation task to consider the possibility that children's knowledge of mappings between number words and approximate numerical magnitudes may have been underestimated in previous studies, and to explore the dependence of that knowledge on children's verbal counting skill. Previous studies claiming that children do not produce larger estimates for sets outside their counting range binned children into only two groups ("skilled" and "unskilled" counters; Lipton & Spelke, 2005). It is possible that this method could lead to underestimation of children's abilities. If some "unskilled counters" tend to produce lower numbers overall while others tend to produce higher numbers, group estimation functions could appear flattened even if individual children did produce larger estimates for sets outside their counting ranges. Also, all of the children in the Lipton and Spelke study succeeded at counting with lower numbers, such that numbers ranging from 20 to 60 were within their counting. Because there are many 5-year-olds who have not yet mastered counting in the 20–60 range, it would be desirable to determine whether or not the previous results hold for these as-yet-untested children.

We first explored the possibility that a similar task might yield a different result if it were administered to participants representing a wider range of counting skill levels, and if a more finely grained counting-level categorization scheme were employed. We presented children with a counting assessment task and partitioned our sample into three counting skill groups. Children and a group of adult participants completed an estimation task adapted from Lipton and Spelke (2005). In Experiment 2, we presented children with a more extensive numerical estimation task, which allowed us to examine individual rather than group performance in order to determine whether individual children would produce larger estimates for larger sets even outside their productive counting ranges.

1. Experiment 1

All participants completed an estimation task, and child participants completed a verbal counting assessment, in order to determine their understanding of the relation between large number words and numerical meanings both within and beyond their productive verbal counting ranges. The estimation task assessed abilities to produce verbal estimates of the number of objects in a visually presented array. Previous research has shown that adults' estimates and children's estimates for numerosities within their counting range are linearly related to the target (presented) numerosity, whereas children's estimates for numerosities outside their counting range are unrelated to the target numerosity (Huntley-Fenner, 2001; Lipton & Spelke, 2005). If knowledge of the relation between number words and nonsymbolic numerosities depends on the ability to count to those words, children should produce linearly increasing estimates only for numerosities within their counting ranges, as reported by Lipton and Spelke (2005). If knowledge of the relation is not dependent on counting skill, however, children of all counting levels might produce linearly increasing estimates, even for numerosities outside of their productive counting range.

Because the Lipton and Spelke study contained no children who were comparable to our least skilled counters (who could not successfully count to 35), our least skilled group is less skilled at counting than any of the children tested previously. If children only know that number words map to approximate magnitudes in an ordered fashion (such that words appearing later in the sequence map to larger sets) for words within the counting range they have already mastered, we should expect the least skilled counters to produce verbal estimates unrelated to the larger target numerosities.

1.1. Method

1.1.1. Participants

All child participants were from central Connecticut families ($n = 50$; mean age 5–10, range 4–10 to 6–7). Adult participants were students in an Introductory Psychology course at Wesleyan University and received course credit. A range of ethnicities and SES levels was represented in the sample; the majority of participants were Caucasian.

1.1.2. Counting assessment

The procedure was modified from Lipton and Spelke (2005). Each child was tested individually in a quiet laboratory setting. Children were asked, "What is the highest number you can count to?" and were then asked to begin counting. Children were stopped if they counted successfully to 35 and were then tested on their ability to complete counting sequences and make correct decade changes. The experimenter began the sequences "55, 56, 57", "78, 79", "95, 96" and "117, 118" and the children were asked to continue counting after the experimenter stopped. If a child was unable to count to 35 without errors, the child was given only the sequence "23, 24, 25." If a child made an error and then self-corrected, the corrected answer was accepted. Children who were unable to count to 35 without error were classified as Level 1 counters, children who made errors between 35 and 60 were classified as Level 2 counters, and children who counted to 60 successfully were classified as Level 3 counters. This group included children who counted perfectly to 120 and children who made errors between 60 and 120. Sixty was chosen as a cutoff point because it was previously shown to be meaningful in assessing children's estimation patterns (Lipton & Spelke, 2005), and 35 was chosen because we encountered many children who could count only to 29.

1.1.3. Estimation task

Forty-nine children and 24 adults participated in this task (one child did not participate due to shyness). The procedure was adapted from Lipton and Spelke (2005). Participants were tested in a quiet laboratory room and were initially presented with a card containing one circle and a card containing 300 circles. The experimenter said, "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 1] and less than this [pointing to card with 300] and ask you to guess how many objects are on each card." Participants were first presented with three practice trials with numerosities under 10 (practice trial cards contained 4, 6, or 7 pink diamonds). Participants then completed 14 test trials with larger numerosities. All participants were presented with the same set of test trial cards containing 20, 40, 60, 80, 100, 120 or 140 pink diamonds (two cards per numerosity). If a participant guessed a number higher than 300, the participant was reminded that all cards contained fewer than 300 objects. To partially control for continuous variables, each participant saw two different sets of cards, one in which object size remained constant and one in which the total summed area of the objects remained constant (such that each participant provided a total of two estimates per numerosity, as in Lipton & Spelke, 2005). Mildly positive feedback was given after all trials (e.g. "Good job").

Data from 17 children were excluded from further analysis due to inattention (one Level 1 counter and one Level 3 counter), difficulty understanding children's word pronunciation (one Level 1), explicitly counting the elements (one Level 3), imaginary estimates (such as "eighty-ninety;" one Level 1, one Level 3), sequential estimates (e.g. 16, 17, 18, 19 for sequential cards; three Level 1s, three Level 2s), or experimenter error (one Level 1, two Level 2s, one Level 3). Data from two adult participants were not included due to experimenter error during the task.

Seven of the children who contributed data to the analyses reported below (three Level 1 counters, one Level 2 counter, and three Level 3 counters) received a slightly modified procedure. In the counting assessment, the children who could count to 35 ($N = 4$ of 7) were allowed to continue counting, rather than being stopped at that point and presented with decade-change sequences. Counting skill level was assessed as before. In the estimation task, these seven children completed the original procedure plus six additional test trials (2 trials each with cards containing 15, 25 or 35 diamonds). This modification was introduced for two reasons. First, we suspected that the decade-change sequences originally used to assess counting skills (adapted from Lipton & Spelke, 2005) could have led some children to produce sequential estimates (and to be excluded as a result). Second, after encountering unexpectedly large numbers of 5-year-olds who could not count to 35, we wished to collect within-count-range and outside-count-range estimation data for this group as well as for more highly skilled counters. Estimates for these seven participants' six additional trials were not included in the analyses.

1.2. Results

1.2.1. Counting assessment

Children demonstrated a wide range of counting abilities and were distributed in a roughly even fashion among the three counting levels. Seventeen children were classified as Level 1 counters (age range 4–10 to 6–1, mean 5–3), 13 children were classified as Level 2 counters (age range 4–10 to 6–7, mean 5–2), and 20 children were classified as Level 3 counters (age range 4–10 to 6–6, mean 5–9). Four Level 3 counters made errors above 60, and 16 made no mistakes. We grouped our counters into three categories rather than the two categories used in previous studies ("unskilled" and "skilled" counters; Lipton & Spelke, 2005). The two-category method would categorize both our Level 1 and Level 2 groups as "unskilled" counters; however, because all of Lipton and Spelke's participants had mastered the count sequence up to about 60, all of our Level 1 counters are in fact *less* skilled than all of their "unskilled" counters. Our Level 2 counters are perhaps most similar to their "unskilled" counters, and our Level 3 counters are similar to their "skilled" counters (though unlike the "skilled" counters, our Level 3 counters did not all perform perfectly: this group includes a few children who made errors between 60 and 120). The Lipton and Spelke study contained no children who were comparable to our Level 1 counters, who had not even mastered verbal counting to 35.

1.2.2. Estimation task

The results of the estimation task are depicted in Fig. 1. Consistent with Lipton and Spelke (2005), we found no difference in estimates for the constant object size versus constant total area sets for adults or children of all counting levels.

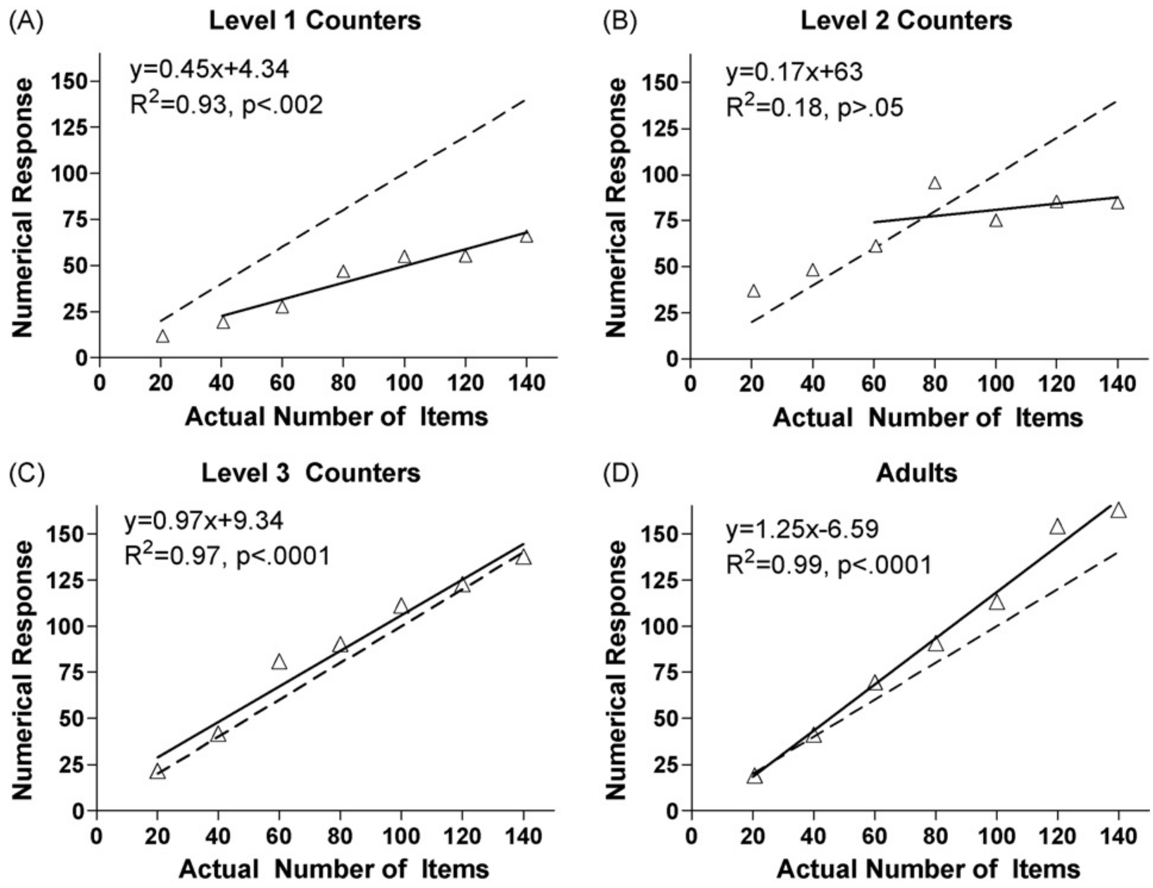

Fig. 1. Estimation performance (triangles) and linear regression lines (solid) for participants in Experiment 1. Dotted lines indicate perfect performance. (A) Level 1 counters: the regression line applies only to estimates produced in response to numerosities outside this group's verbal counting range. (B) Level 2 counters: the regression line applies only to estimates produced in response to numerosities outside this group's verbal counting range. (C) Level 3 counters: most or all of the presented numerosities were inside this group's verbal counting range. (D) Adults.

1.2.3. Level 1 counters

For Level 1 counters ($n = 9$, mean age 5–2), all or nearly all of the presented numerosities were outside their productive counting range. Unsurprisingly, Level 1 counters, most of whom do not know the larger counting words at all, generally produced estimates that were smaller than the target numerosities. However, their estimates still increased linearly in proportion to the target numerosity ($y = 0.46x + 4.17$, $R^2 = .96$, $p < .001$). Because some (although not all) Level 1 counters produced verbal count lists up to 20, we considered the data with estimates of 20-item sets excluded as well. The same result was obtained when 20-item sets were excluded ($y = 0.45x + 4.34$, $R^2 = .93$, $p < .002$). Level 1 counters, although they could not produce a verbal counting list up to 35, did produce significantly larger estimates for larger sets.

1.2.4. Level 2 counters

The Level 2 counters ($n = 8$, mean age 5–7) also produced estimates that were linearly related to the target numerosity overall, although many of the target numerosities again fell outside their counting range ($y = 0.40x + 37.71$, $R^2 = .68$, $p < .05$). However, when only the sets outside their productive counting range were considered (sets containing 60–140 items), Level 2 counters did not produce larger estimates for larger sets: the slope of the best-fit line for these estimates was not different from zero ($y = 0.17x + 63$, $R^2 = .18$, $p > .05$).

1.2.5. Level 3 counters

Level 3 counters ($n = 15$, mean age 5–10), whose productive counting ranges included most or all of the presented numerosities, produced remarkably accurate estimates that increased linearly with the target numerosity ($y = 0.97x + 9.34$, $R^2 = .97$, $p < .0001$), as did adults ($y = 1.25x - 6.59$, $R^2 = .99$, $p < .0001$), in accord with previous findings (Lipton & Spelke, 2005).

1.2.6. Tests for potential explanations of less-skilled counters' performance

The analysis described above does not include data from children who produced sequential estimates (e.g. "16, 17, 18, 19" for sequential cards) because it is unlikely that these estimates are representative of children's actual beliefs about the number of items on the cards. Three Level 1 counters and three Level 2 counters were therefore excluded from the original analysis (no Level 3 counters produced sequential estimates). Because the cards were not presented in order of magnitude, inclusion of data from children who produced sequential estimates would increase the likelihood that mean estimates would appear unrelated to the presented numerosities. Although Lipton and Spelke (2005) do not report that any participants gave sequential estimates, it is possible that their study did include data from children who gave sequential estimates, and that the inclusion of these data could be partially responsible for the differences in our findings. We therefore analyzed the present data with these six children included, with the same result: Level 2 counters still did not produce increasing estimates for larger sets that fell outside of their productive count range, and Level 1 counters still provided estimates that increased linearly in proportion to the target numerosity, whether or not 20-item sets were included in the analysis (with 20-item sets: $y = 0.35x + 7.47$, $R^2 = .94$, $p < .0005$; without: $y = 0.35x + 7.9$, $R^2 = .91$, $p < .004$).

To rule out two other possible explanations of our data, we conducted two additional tests. First, it is possible that the use of means as a measure of the estimates' central tendency allowed outliers to exert a large influence on our results. We recalculated the estimates using medians instead of means, with the same outcome. Most importantly, the median estimates for the Level 1 counters were larger for larger presented numerosities, even though all or nearly all estimates involved numerosities outside the productive counting ranges of these children. This was the case whether or not sets with 20 elements (which are within the productive counting range of some Level 1 counters) were included in the analysis (with 20-item sets: $y = 0.51x + 9.4$, $R^2 = .88$, $p < .0005$; without: $y = 0.56x - 9.7$, $R^2 = .86$, $p < .01$). Second, it is possible that relatively unskilled counters, like our Level 1 children, are simply more likely to produce large round number estimates like "100" for very large numerosities. We recalculated the estimates excluding children who produced the estimates "100" or "200" on six or more trials (except when the presented numerosity was indeed 100). This conservative test excludes estimates of "100" even when these estimates are reasonable (such as for presented numerosities of 80 or 120). This method excluded six children from the Level 3 counting group, but the remaining Level 3 counters still produced linearly increasing estimates ($n = 9$, $y = 0.8815x - 1.729$, $R^2 = .9923$, $p < .0001$). Two Level 2 counters were excluded, and those remaining ($n = 6$) did not produce increasing estimates outside their count range (the same result observed for Level 2 counters overall). However, no Level 1 counters were excluded by this method, showing that Level 1 counters were not especially likely to produce large round numbers when presented with large sets. Level 1 counters' linearly increasing estimates for numerosities outside of their count range, therefore, could not be due to repeated production of the estimates "100" or "200."

1.3. Discussion

Overall, the most skilled counters in this study, the Level 3 counters, produced verbal estimates that increased in a linear fashion with set size, as expected: most or all of the numerosities presented fell within the verbal counting range that these children had mastered. Level 3 counters' estimates, like those of adults, were relatively accurate (on average), suggesting that these children had formed a mapping between the verbal count list and the approximate numerical representations supplied by the nonverbal analog magnitude system. This finding replicates the performance of Lipton and Spelke's (2005) "skilled" counters. The current experiment also included a set of slightly less-skilled counters, the Level 2 counters. These children were somewhat comparable to the previous study's "unskilled" group in verbal counting ability. Level 2 counters as a group did not produce larger estimates for larger sets for numerosities outside their productive verbal counting range, again replicating earlier findings with children at roughly the same level of counting skill (Lipton & Spelke, 2005).

Our results deviate from those of earlier experiments in one important way. The previous study did not include any children who were unable to count to 35 (Lipton & Spelke, 2005), while the current study did include such a group, the Level 1 counters. In this experiment, Level 1 counters were asked to produce verbal estimates for sets containing numerosities that always or nearly always fell outside of their productive verbal counting range. Unsurprisingly, Level 1 counters did not produce accurate estimates, as they did not yet know the number words in the relevant portion of the count list. The Level 1 counters, nevertheless, *did* produce significantly larger verbal estimates for larger sets, and this finding could not simply be attributed to the frequent production of large round numbers like 100 or 200. This pattern of performance suggests that even these very inexperienced counters do have some knowledge of mappings between approximate numerosities and number words, even when the presented sets are associated with number words that fall far beyond their productive counting range. Experiment 2 explores this possibility further.

2. Experiment 2

The results of Experiment 1 suggest that even children who cannot reliably produce a verbal count sequence up to 35 can produce increasing estimates for larger numbers beyond their productive counting range. Because relatively few data points were collected for each individual child, the analyses of Experiment 1 were carried out on group data as in previous work (Lipton & Spelke, 2005). But when relatively unskilled counters are presented with tasks of this sort, there may be a great deal of noise in children's individual response patterns. Perhaps some children happen to produce large estimates when presented with large sets, while other children happen to produce small estimates for larger sets. On average, such noisy response patterns should lead to a group estimation function with a slope not significantly different from zero (as produced by previous "unskilled" counters; Lipton & Spelke, 2005).

It is possible that the Level 1 counters of Experiment 1 appeared to produce estimates that increased with set size only by chance, and that our sample simply happened to contain an unusually large number of children who produced large estimates for large sets. If so, the results of Experiment 1 should not be taken to indicate that Level 1 counters can systematically produce larger estimates for larger sets. To test this possibility, we presented a modified estimation task to a new group of children in order to examine individual estimation patterns. If very inexperienced counters do have some knowledge of mappings between approximate numerosities and number words, the least skilled counters should again produce increasing estimates for larger sets, and there should exist some children in this group whose data exhibit this pattern even at the individual level. In order to provide more information about the variability of individual children's estimates, Experiment 2 required children to produce a larger number of estimates for each presented numerosity.

2.1. Method

2.1.1. Participants

Thirty-two children from central Connecticut families completed the study (mean age 5–7, range 4–10 to 6–8). A range of ethnicities and SES levels was represented, and the majority of participants were Caucasian.

2.1.2. Procedure

Children were tested in a quiet laboratory room. The estimation task was presented first, followed by the counting assessment, so that experimenters were unaware of each child's counting skill level during the estimation trials. Four participants completed the counting assessment before the estimation task. In the estimation task, the experimenter presented the child with a piece of paper covered with stickers and asked the child to guess how many stickers were on the paper. The estimation stimuli consisted of letter-sized pieces of paper that had been covered with stickers and laminated. All stickers on a piece of paper were alike, but stickers of different shapes (hearts, tennis balls, smiley faces, footballs, basketballs, stars) and sizes (long axes or diameters approximately 1.4 cm, 1 cm, 1 cm, 1 cm, 1 cm, and 0.5 cm, respectively) were used to maintain interest and vary non-numerical stimulus properties. The 56 test stimuli included 8 stimuli containing fewer than 20 stickers (not analyzed) and 48 test stimuli containing 20, 35, 60, or 100 stickers (12 for each numerosity).

The experimenter began by showing the child a paper with 1 circular blue sticker and a paper with 200 circular blue stickers. The experimenter said, “This paper has 1 sticker on it, and this paper has 200 stickers on it. That’s too many for you to count, but there are 200 stickers on there. I’m going to show you more papers and ask you to guess how many stickers are on each paper. All of the papers I’m going to show you have more than this [pointing to the paper with 1 sticker] and fewer than this [pointing to the paper with 200 stickers]. Now I know you know how to count, but I want you to just *guess* how many stickers are on each piece of paper.” The child was then shown 56 stimuli, one at a time, and asked for an estimate for each one. Participants who attempted to count the stickers were reminded that this was a guessing game and were asked to make a guess. Participants who guessed a number above 200 were reminded that all of the papers had fewer than 200 stickers, and were asked to make a new guess. The upper limit was changed to 200 instead of 300 (as it was in Experiment 1) in case the introduction of a number as large as 300 had influenced, and perhaps inflated, children’s estimates.

After completing the estimation task, children were asked to count as high as they could. If children corrected themselves after making an error, the correction was accepted. They were stopped after one uncorrected error or when they reached 100. Children were classified by counting skill level as in Experiment 1.

2.2. Results

Four children (two Level 1 counters, one Level 2 counter, and one Level 3 counter) were excluded from further analysis due to extreme inattention, sequential guesses, or imaginary guesses (e.g. “forty-seventy”). Twenty-eight participants were included in the analyses. Eight were classified as Level 1 counters (mean age 5–2, range 4–10 to 5–4), six were classified as Level 2 counters (mean age 5–7, range 5–0 to 6–2), and 14 were classified as Level 3 counters (mean age 5–10, range 5–1 to 6–8).

2.2.1. Level 1 counters

For Level 1 counters, all or nearly all of the presented sets fell outside the productive verbal count range. One Level 1 counter produced a verbal count list up to 29, and a second produced a list of length 20. The others made mistakes below 20. As in Experiment 1, Level 1 counters as a group produced estimates smaller than the presented numerosities, but their estimates did increase linearly with the presented numerosity ($y = 0.43x + 11.12$, $R^2 = .94$, $p < .03$). Even if we consider only the largest sets presented, 60 and 100, Level 1 counters produced larger estimates for larger sets: Estimates produced for sets of 100 were significantly larger than estimates produced for sets of 60 ($t(89) = 2.14$, $p < .02$). Individual estimates from the eight Level 1 counters are depicted in Fig. 2. Although a few of these children did not produce larger estimates for larger presented numerosities outside their count ranges ($n = 3$, all $F < 1.5$, $p > .05$), when their data were analyzed individually the majority of the Level 1 counters did produce significantly larger estimates for larger sets outside their count ranges (that is, the slopes of the best-fit lines for these individual children’s estimates outside their productive counting ranges were greater than 0; $N = 5$, all $F > 13$, $p < .002$).

2.2.2. Potential strategies used by Level 1 counters

Did the Level 1 counters in fact produce their estimates by relating number words to their internal system of ordered mental magnitude representations? In order to produce the observed data, children must have assessed the approximate numerical size of the sets, and the mental magnitude system must be implicated in this process. It is not necessarily the case, however, that children produced each estimate by retrieving a large number word that s/he had previously mapped to a corresponding mental magnitude. There are at least two different ways in which these unskilled counters might produce larger estimates for larger sets when presented with an estimation task involving numerosities outside their stable verbal count range. When presented with an estimation task, some children might map

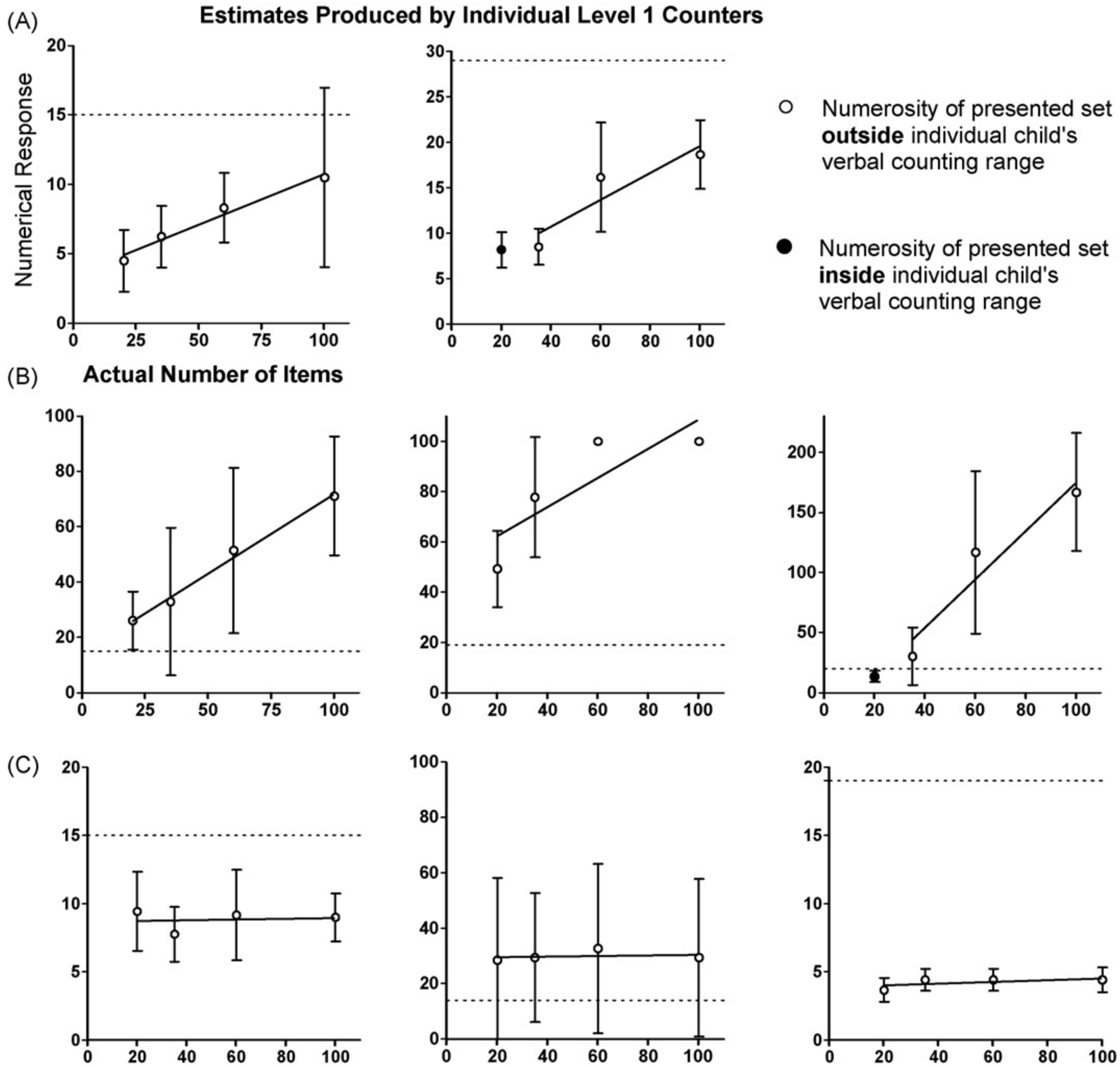

Fig. 2. Individual data from eight of the least skilled counters (Level 1 counters) in Experiment 2. Horizontal dotted lines indicate the limits of each child's productive verbal count list. (A) Children whose data are depicted in row A produced larger estimates for larger sets, and generally did so by producing number words from within their stable verbal counting ranges. (B) Data from children who gave larger estimates from larger sets by producing number words outside their stable counting range. (C) Data from children who did not produce larger estimates for larger sets; the slopes of the best-fit lines for these children's estimates did not differ significantly from zero. No child produced estimates that decreased with increasing set size.

the smallest presented sets onto the lower end of their known count list and map the largest presented sets onto the higher end of their known count list. This strategy would lead to underestimation, but it could lead children to produce significantly larger estimates for larger presented numerosities. Other children might know a few large number words beyond the count sequence they could stably produce, such that a child who could only count to 29 might frequently produce estimates of "seventy-seven" or "eighty" for the largest sets presented, for example, and estimates of "fifty" for the next-largest sets. The Level 1 counters who did produce larger estimates for larger sets showed some evidence of each of these strategies: two of these children generally produced estimates from within their stable verbal counting range, and three of them produced numbers that were not part of the sequence they could stably produce (Fig. 2).

It is also possible that children could produce linearly increasing estimates on average, even if they only formed very limited mappings between number word-magnitude pairs at the larger end of the range. For example, consider a hypothetical child who counts reliably only to 29, but who knows that the word "eighty" refers to a large number. This child might produce the word "eighty" very frequently when presented with very large sets and less frequently when presented with smaller sets. Such a pattern of performance would lead the child to produce estimates that increased, on average, with the numerical size of the presented set, but it would not necessarily mean that the child knew about the relation between number words and numerical magnitudes in general. A mapping of sorts between number words and mental magnitudes would still be involved in this strategy, but such a mapping would be limited to just a few word-magnitude pairs.

Qualitatively, children's estimates did not fit this pattern. Children produced many different number words, not just a few words, even for very large sets. An analysis of the variability of children's estimates offers a more quantitative assessment. A child using the strategy described above would frequently produce the same large estimates as the size of the presented set increased, such that the COV (the coefficient of variation, or the ratio of the standard deviation to the mean), would decrease with increasing set size. One of the well-known signatures of estimation based on mappings to the mental magnitude system, on the other hand, is a constant COV: the standard deviation of the estimates increases in proportion to the mean (Gallistel & Gelman, 1992, 2000; Huntley-Fenner, 2001; Whalen et al., 1999). For example, when 5-year-olds are asked to estimate the number of items in rapidly presented sets of 5, 7, 9, or 11 items, COV scores do not increase with set size (Huntley-Fenner, 2001). Therefore, if children's COV scores in the present study decrease with set size, the data are consistent with the endpoint-based guessing strategy described above. If children's COV scores remain constant, the data are consistent with the idea that children's estimates are not based simply on guesses anchored to a few extreme set sizes, but rather are based on broader mappings to mental magnitudes. The COV scores of the Level 1 counters remained constant with increasing set sizes, $F(3,21) = 0.8514$, $p > .05$, providing evidence against the endpoint-based guessing strategy and suggesting that children did provide estimates based on broader mappings between number words and mental magnitudes.

2.2.3. Level 2 counters

One Level 2 counter successfully produced a verbal count sequence up to 35; two Level 2 counters counted up to 39, and three produced a count sequence up to 59. Estimates produced for sets of 100 were significantly larger than estimates produced for sets of 60 (the two presented numerosities that were outside the verbal count ranges for these children; $t(71) = 2.13$, $p < .05$). When estimates across the entire range of presented numerosities were examined at the individual level, three of the six did produce significantly larger estimates for larger sets overall (all $F > 5$, $p < .03$); two of these children had counted successfully to 59, and one had counted to 35. Three of the six Level 2 counters did not produce larger estimates for larger presented numerosities (all $F < 4$, $p > .05$). Level 2 counters as a group produced estimates that did not increase as a linear function of the target numerosity, even though half of the presented numerosities in Experiment 2 fell inside their verbal counting range. There was no effect of set size on COV for the Level 2 counters, $F(3,15) = 0.855$, $p > .05$.

2.2.4. Level 3 counters

The Level 3 counters, whose productive verbal count ranges included all of the presented numerosities, produced accurate group mean estimates that increased linearly with the target numerosity ($y = 0.86x + 6.99$, $R^2 = .97$, $p < .02$). All 13 Level 3 counters, considered individually, produced larger estimates for larger numerosities (all $F > 5$, $p < .03$). There was no effect of set size on COV for Level 3 counters, $F(3,36) = 1.759$, $p > .05$.

2.3. Discussion

These results confirm and extend the findings of Experiment 1, showing that children are able to produce larger number words for larger sets, even when those sets are outside of their individually-determined productive counting ranges. Five of eight of the least skilled Level 1 counters produced significantly larger estimates for larger presented sets, even though those sets contained numerosities corresponding to numerals that fell well outside the range they could reliably recite. The variability observed in children's estimates was not consistent with the idea that children simply knew a few large number words, producing these words more frequently when presented with especially large sets and less frequently when presented with smaller sets. Rather, even the least skilled counters showed evidence of partial mappings between number words and mental magnitudes. Children evidently can produce larger number words for larger sets, even before they have learned the portion of the verbal count list that would be required to produce accurate estimates.

The group data for Experiment 2 echo the results of Experiment 1 in three ways. First, the most skilled (Level 3) counters produced relatively accurate estimates that increased linearly with the numerosity of the presented set. Second, the children who were the least skilled at producing a verbal count list (Level 1 counters) also produced estimates that increased linearly with set size, although, presumably because of their limited command of the verbal count list, their estimates were quite low relative to the actual numerosities presented. Third, the children who demonstrated intermediate levels of verbal counting skill (Level 2 counters) did not reliably produce larger estimates for larger sets (these children produced significantly larger verbal estimates for sets of 100 than for sets of 60 in Experiment 2, but when their estimates for all set sizes were considered, they did not produce significantly larger estimates for larger sets on average). These performance patterns and their relation to earlier findings are discussed below.

3. General discussion

We investigated the relationship between children's productive verbal counting skill and their understanding of mappings between number words and numerical magnitudes in the large number range. Five-year-old children were asked to estimate the number of items on a card and were sorted into three groups based on counting skill (Level 1, Level 2, and Level 3 counters). The main finding, supported by two experiments, was the novel result that even children who have mastered very little of the verbal count list demonstrate some knowledge of how approximate numerical magnitudes map onto large number words. This understanding holds not only for the range of numerosities corresponding to number words that the child can reliably produce when asked to count verbally, but also for those numerosities well beyond the verbal counting range.

The most skilled counters in the present experiments (Level 3 counters, who could produce a verbal count list at least up to 60 and usually up to 100 or beyond) offered verbal estimates that increased in a linear fashion with the numerosity of the presented set. Estimates produced by Level 3 counters were quite accurate on average. All or nearly all of the numerosities presented corresponded to number words within the productive verbal counting range of these children. This finding replicates previous results (Lipton & Spelke, 2005). Children whose verbal counting skill was classified as intermediate (Level 2 counters, who counted correctly at least to 35 but erred before 60) did not reliably produce larger estimates for larger sets as a group, although some individual children did so. The least skilled counters taken together (Level 1 counters, none of whom could count reliably to 35 and many of whom failed to count even to 20) reliably produced verbal estimates that increased linearly in proportion to the presented numerosities, even though most or all of the numerosities presented corresponded to number words that did not fall within these children's productive verbal count ranges. Furthermore, this pattern of performance was identifiable not only in the group data of Experiments 1 and 2, but also in the individual data produced in Experiment 2 by these least skilled counters.

We interpret these findings to mean that children can produce meaningful estimates even for sets that they have not yet learned to count. Children who haven't yet mastered the verbal count sequence from 1 to 100 (or even from 1 to 35) nevertheless possess a partial understanding of the mapping between approximate numerical quantities and large number words. We also report two novel (though preliminary) findings from the data produced by the least skilled counters in this study: some of these children flexibly extended their limited verbal count list to map onto the entire range of presented sets, while others labeled presented sets with number words that were far beyond the highest word in the verbal count list they could reliably produce. Recent work has shown that mappings between numerical labels and numerosities are surprisingly flexible even in adults (Izard & Dehaene, 2008). Additional studies of relatively unskilled counters' estimation performance in similar tasks may allow us to determine the prevalence of these flexible mapping patterns in children, and to identify the specific strategies that underlie the mapping processes involved.

The present conclusions are at odds with the conclusions of some previous studies (Lipton & Spelke, 2005), but the data themselves may not be incompatible. This is because our Level 3 counters were most comparable to Lipton and Spelke's "skilled" counters, and both groups produced comparable data. The present Level 2 counters were most comparable (though not identical) to their "unskilled" counters, and they also produced comparable data. Previous research using this task did not test children with the verbal counting skills of our Level 1 counters. The inclusion of this least skilled group in the present study, along with the current finding that these least skilled counters do produce larger estimates for larger sets, may account for these conflicting interpretations, although it does not provide an explanation of our results.

3.1. Possible explanations of intermediately-skilled counters' inconsistent performance

The inconsistent performance of Level 2 counters, while comparable to that found by Lipton and Spelke (2005), was surprising given that the Level 1 counters produced larger estimates for larger sets even though they had a shorter productive count list than the Level 2 counters. One possible explanation of our findings is that a larger group of Level 2 counters would, on the whole, produce larger estimates for larger sets. Although it is ideal to keep experimenters blind to individual children's counting ability, this approach necessarily results in a relatively small number of counters in the two less-skilled categories, especially in the Level 2 group. (We did not test younger 4-year-olds both because of the length of the estimation session and because such children may not yet have begun to map number words to mental magnitudes; LeCorre & Carey, 2007.) Previous studies with larger numbers of children fairly similar to our Level 2 counters, however, have found similar patterns of performance (Lipton & Spelke, 2005), suggesting that a larger sample would not be likely to yield a different result.

A different possible explanation of this pattern might be strategic variability among the Level 2 counters. Siegler and colleagues (Opfer & Siegler, 2007; Siegler, 2007) have proposed that periods of conceptual, representational, or technique-related instability may be necessary in order to learn effectively, as they allow children to experiment with a variety of strategies. Possibly the Level 2 counters' performance is evidence of a period of instability that occurs once children have mastered part of the count list. If so, Level 2 counters might make use of a variety of strategies, leading to inconsistent performance on the estimation task overall.

Some indirect support for this idea may come from previous number-line estimation work. In one such study (Booth & Siegler, 2006), second graders appeared to use one strategy when making estimates of numbers within their familiar range (1–100) and another strategy in their unfamiliar range (1–1000). In a more recent study (Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008), 5-year-olds' number-line estimation performance was best fit by a segmented linear model with a steep positive slope for smaller numbers and a shallow slope for larger numbers. The authors suggested that the two lines correspond to estimates made for numbers within versus outside the child's familiar count list, demonstrating a flexibility in strategy use during the development of awareness of the count sequence.

Similarly, our Level 2 counters could be recruiting different strategies within versus outside their familiar count range. For example, Level 2 counters might possess sufficient knowledge of the proper mappings between number words and analog magnitudes to realize that sets of 20 or 35 elements are within the range they have mastered, and might therefore attempt to produce relatively accurate estimates for such sets. For much larger sets outside their counting range, Level 2 counters might reject such a strategy and resort instead to a different one – something like "just say one of the few very big number words I know." There may be some evidence of such a strategy among our Level 2 counters (see the discussion of logarithmic vs. linear models of performance, below), but due to the relatively small number of Level 2 counters identified here, further data are necessary to explore this possibility. Also, the finding that the Level 2 counters in Experiment 2 did produce larger estimates for sets of 100 elements than for sets of 60 elements suggests this group's strategy, when faced with a very large set, was not simply "say one of the few very big number words I know." Future studies aimed specifically

Table 1
Comparison of linear and logarithmic fits for the median estimates of Level 1, Level 2, and Level 3 counters in Experiments 1 and 2 (for method see Opfer, n.d.).

Exp.	Counting skill level	lin R^2 (medians)	log R^2 (medians)	p-Value	log R^2 > lin R^2
1	1	0.8795	0.8037	$p > .05$	
1	2	0.6289	0.7525	$p > .05$	
1	3	0.9231	0.8825	$p > .05$	
2	1	0.6714	0.8605	$p > .05$	
2	2	0.9384	0.9987	$p < .005$	yes
2	3	0.9742	0.9788	$p > .05$	

Table 2
Comparison of linear and logarithmic fits for the mean estimates of Level 1, Level 2, and Level 3 counters in Experiments 1 and 2 (for method see Opfer, n.d.).

Exp.	Counting skill level	lin R^2 (means)	log R^2 (means)	p-Value	log R^2 > lin R^2
1	1	0.9577	0.9154	$p > .05$	
1	2	0.6810	0.7823	$p > .05$	
1	3	0.9664	0.9682	$p > .05$	
2	1	0.9412	0.9694	$p > .05$	
2	2	0.8869	0.9791	$p < .05$	yes
2	3	0.9734	0.9685	$p > .05$	

at the analysis of individual Level 1 and Level 2 counters' patterns of performance may provide further insight into this unexpected finding.

3.2. Potential relation to number-line estimation studies

Some previous researchers have used a very different type of estimation task, number-line estimation, to argue that estimation accuracy is related to the type of number representation employed (Siegler & Booth, 2004; Siegler & Opfer, 2003). In these studies, an experimenter specifies a number and asks a child to indicate the correct location for that number on a number line. In such tasks, younger children produce the least accurate estimates, and their responses appear to reflect a logarithmic relationship between symbolic number and magnitude. Older children and adults give more accurate estimates that appear to reflect a linear representation of number. For example, 5-year-olds' responses for a 1–100 number line tend to be fit best by a logarithmic function, while 7-year-olds' responses for the same number line appear linear (Siegler & Opfer, 2003; but see Ebersbach et al., 2008, for a different view). Siegler and colleagues suggest that performance on these tasks indicates that children's internal representations of number are logarithmic early in life, shifting to become linear as they gain experience with the mapping between number words and magnitudes.

In the present estimation tasks, we did not find evidence to support this view. Most importantly, in neither experiment did the least skilled (Level 1) counters produce estimates that were best fit by a logarithmic curve. In Experiment 1, logarithmic regression did not provide a significantly better fit than linear regression for any group (see Table 1 for median estimates and Table 2 for mean estimates). In Experiment 2, a logarithmic fit was significantly better than a linear fit only for the moderately skilled Level 2 counters' estimates (see Tables 1 and 2). For two reasons, however, we suggest that this finding should not be taken to imply that the Level 2 counters of Experiment 2 possess a logarithmically organized internal representation of number.

First, the task of Experiment 2 elicited many estimates per child for every set size tested, allowing us to look closely at individual children's estimates. No individual child's estimates were better fit by logarithmic regression than by linear. The finding of a significantly better logarithmic fit for the Level 2 group in Experiment 2 appears to arise from individual Level 2 counters who did not produce larger estimates for larger sets outside their counting range; moreover, a few of these counters repeated the same large numbers for nearly every large set presented. For example, one Level 2 counter provided an estimate of "90" for every set of 60 or 100 items, except for two presentations of 60 items which were labeled "80." In our view, this kind of estimation pattern implies the recruitment of a sensible strategy ("use the few large number words you know whenever you see a very large set") rather than the presence of an internal representation of number that is logarithmically organized.

The second reason we suggest that this finding should not be taken to imply that the Level 2 counters of Experiment 2 possess a logarithmic representation of number is as follows. Although both linear and logarithmic functions provided good fits for these children's data (see Tables 1 and 2), accepting the best-fitting logarithmic model as the best explanation of this group's data introduces a serious problem. The logarithmic fit predicts that children should guess that sets of 10 or fewer elements actually contain zero elements, or a negative number of elements. This is both intuitively unlikely and incompatible with children's responses for the small sets presented during the experiments.

Although our results (like those of Lipton & Spelke, 2005) do not appear to support the notion of a logarithmic-to-linear shift in children's numerical estimation, the present task differed from number-line estimation tasks in many ways. It is possible that the number-line task is especially likely to elicit patterns of responses indicative of a logarithmic-to-linear shift and that these results may not generalize to other types of estimation tasks (Noel, Rousselle, & Mussolin, 2005; but see Booth & Siegler, 2006). It is also important to note that because the present experiments were not designed as a test of the log-to-linear shift hypothesis, we did not sample heavily in the small number range for the purpose of distinguishing logarithmic from linear estimation functions, as did some previous researchers (Siegler & Booth, 2004). Finally, recent work suggests that a segmented linear model may provide a better explanation of children's number-line estimation performance than does a logarithmic model (Ebersbach et al., 2008); it is possible that further experiments specifically designed to adjudicate among these competing models may reveal similar findings for large-set estimation tasks like those employed in the present study.

3.3. Conclusions

Many studies continue to explore the processes underlying the initial acquisition of number word meaning (LeCorre & Carey, 2007; Sarnecka & Carey, 2008) and the changes in children's mappings of number words to magnitudes that occur throughout elementary school (Opfer & Siegler, 2007). Fewer studies have looked at the processes underlying the initial acquisition of mappings between larger number words and the approximate numerosities of the nonverbal mental magnitude system. Therefore, there is much to be learned about the way in which children come to form these mappings and about the extent of children's knowledge of the properties of number words beyond their verbal count lists.

Preschool children who have not yet mastered verbal counting do understand that number words both within and beyond their counting range refer to specific, exact numerosities (Condry & Spelke, 2008; Lipton & Spelke, 2006; Sarnecka & Gelman, 2004; Wynn, 1992), but previous research has suggested that such children do not have the ability to produce larger verbal estimates for larger sets (Lipton & Spelke, 2005). Here, we show that children apparently do know that larger numbers map to approximate magnitudes in an ordered fashion, such that words that appear later in the sequence map to larger sets, for numbers both within and beyond their counting range. These findings provide evidence that children possess at least a partial knowledge of this mapping before they have integrated many number words into their count sequence, and that mastering the count sequence does not guide children's acquisition of this aspect of number word logic. Further research may be able to provide insight into the specific mechanisms underlying children's formation of mappings between number words and approximate numerosities.

Acknowledgments

The authors thank AnjaLi Carrasco, Rachel Jacobson, Jessica Tsai, Annie Paladino, Julie Neuspiel, and Elise Herrig for help with data collection; Elizabeth Spelke, Susan Carey, and anonymous reviewers for comments on previous versions of the manuscript; and the parents, children, and preschool coordinators who made this work possible. This project was supported by Hughes Summer Research Fellowships to A.S. and J.S. and a Wesleyan University Project Grant to H.B.

References

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. S. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*, 14116–14121.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *41*, 189–201.

Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind & Language*, *16*, 37–55.

Carey, S. (2004). Bootstrapping & the origin of concepts. *Dcedalus*, *Winter*, 59–68.

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural numbers. *Journal of Experimental Psychology: General*, *137*, 22–38.

Cordes, S., & Brannon, E. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*, 803–808.

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

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

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year-old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology*, *99*, 1–17.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Frank, M., Everett, D., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Piraha language and cognition. *Cognition*, *108*, 819–824.

Gallistel, C. R. (2007). Commentary on Le Corre & Carey. *Cognition*, *105*, 439–445.

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

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

Gilmore, C. K., McCarthy, S., & Spelke, E. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*, 589–592.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496–499.

Hauser, M., & Spelke, E. (2004). *Evolutionary and development foundations of human knowledge: A case study of mathematics*. *The cognitive neurosciences* Cambridge: MIT 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.

Izard, V., & Dehaene, S. (2008). Calibrating the number line. *Cognition*, *106*, 1221–1247.

LeCorre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395–438.

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

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

Lipton, J., & Spelke, E. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, *76*, 978–988.

Lipton, J., & Spelke, E. (2006). Preschool children master the logic of number word meanings. *Cognition*, *98*, B57–B66.

Moyer, R., & Landauer, T. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Noel, M.-P., Rousselle, L., & Mussolin, C. (2005). Magnitude representation in children: Its development and dysfunction. In J. Campbell (Ed.), *Handbook of mathematical cognition*. New York: Psychology Press.

Opfer, J. (n.d.). *Analyzing the number-line task: A tutorial*. Retrieved on March 3rd, from http://www.psy.cmu.edu/~siegler/publications-all.html.

Opfer, J., & Siegler, S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*, 169–195.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499–503.

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

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*, 662–674.

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, *92*, 329–352.

Siegler, S. (2007). Cognitive variability. *Developmental Science*, *10*, 104–109.

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

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

Spelke, E. S., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition*, *78*, 45–88.

Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: advances in the investigation of language and thought*. Cambridge, MA: MIT Press.

Whalen, C., Gallistel, G., & 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.
Wynn, K. (1995). Origins of numerical knowledge. *Mathematical Cognition*, *1*, 35–60.
Wynn, K. (1998). Psychological foundations of number: Numerical competence in human infants. *Trends in Cognitive Sciences*, *2*, 296–303.
Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representation. *Cognition*, *89*, B15–B25.
Xu, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.