CHILD DEVELOPMENT

Child Development, May/June 2014, Volume 85, Number 3, Pages 1306–1319

Young Children's Interpretation of Multidigit Number Names: From Emerging Competence to Mastery

Kelly S. Mix
Michigan State University

Richard W. Prather and Linda B. Smith
Indiana University

Jerri DaSha Stockton
Michigan State University

This study assessed whether a sample of two hundred seven 3- to 7-year-olds could interpret multidigit numerals using simple identification and comparison tasks. Contrary to the view that young children do not understand place value, even 3-year-olds demonstrated some competence on these tasks. Ceiling was reached by first grade. When training was provided, there were significant gains, suggesting that children can improve their partial understandings with input. Findings add to what is known about the processes of symbolic development and the incidental learning that occurs prior to schooling, as well as specifying more precisely what place value misconceptions remain as children enter the educational system.

Research has established that most children enter school with basic counting skills and a firm conceptual understanding of numbers up to 10. From an early age, children can name, match, order, and calculate with these quantities (for a review, see Mix, Huttenlocher, & Levine, 2002). This early competence provides a strong foundation upon which to build conventional skills and most children fare well as long as single-digit number facts are involved. However, when multidigit numbers are introduced, even competent children struggle—struggles that manifest themselves not only in weak place value concepts, but also in rote, error-prone application of algorithms for multidigit calculation (Fuson, 1990; Kamii, 1986; Kouba et al., 1988; Labinowicz, 1985; Miura, 1987; Ross, 1990; Towse & Saxton, 1997). This pattern is concerning because place value is the gateway to conceptualizing large quantities and more complicated mathematical operations, such as addition with carrying. Moreover, there is a significant relation between children's place value skills in early elementary grades and subsequent problem-solving ability. In short, children who fail to master place value face chronic low achievement in mathematics (Ho & Cheng, 1997; Moeller, Martignon, Wessolowski, Engel, & Nuerk, 2011).

These difficulties have led researchers to conclude that place value notation is fundamentally inaccessible to young children. Some have argued that the spoken and written numeration systems are so different that children fail to see how they are related without specialized instruction, such as lessons using base-10 blocks (e.g., Fuson, 1990; Fuson & Briars, 1990). Others have argued that children lack the logical capacity to comprehend place value notation (Chandler & Kamii, 2009; Fosnot & Dolk, 2001). On either account, it is assumed place value notation is incomprehensible to children without significant development and direct instruction.

However, these assumptions are inconsistent with what we know about early cognition and learning. We know, for example, that toddlers acquire the complex grammatical structures in their native languages simply through exposure—without the need for direct instruction (Aslin & Newport, 2012; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; MacNamara, 1972). The same is true for word learning. Even in complex scenes with numerous potential referents (Smith & Yu, 2008), or long speech streams without gaps between words

This study was supported by a grant from the Institute of Education Sciences (R335A080287) to the first and third authors.

Correspondence concerning this article should be addressed to Kelly S. Mix, College of Education, Michigan State University, 620 Farm Lane, East Lansing, MI 48824. Electronic mail may be sent to kmix@msu.edu.

 0009-3920/2014/8503-0034
DOI: 10.1111/cdev.12197

(Graf-Estes, Evans, Alibali, & Saffran, 2007), infants can segment and learn the meaning of new words with only brief exposure. Researchers have explained these findings in terms of statistical learning—the idea that learners actively make sense of the perceptual stream by tabulating the statistical patterns in it (e.g., Kidd, 2012; Yu, Ballard, & Aslin, 2005). So, for example, infants identify word segments in an uninterrupted speech stream by noticing that /pa/ follows /ba/ more frequently than it follows /do/.

It is possible children acquire a partial understanding of multidigit numerals the same way. Parents and teachers provide relatively little direct input related to number (Gunderson & Levine, 2011; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006) so it is unlikely they deliberately teach children to read multidigit numbers. However, multidigit numerals are ubiquitous in children's environments—as room numbers, phone numbers, and street addresses; in books, calendars, and menus; and throughout stores on packaging, price tags, and signs. Because statistical learning happens rapidly, this admittedly limited exposure could be sufficient to begin extracting certain structural patterns. For example, two-digit number names almost always have the sound /ee/in the middle and three-digit number names almost always have the word *hundred* in them. Obviously, this correlation is not perfect and it breaks down further when longer number names are considered, but it is consistent enough to help children guess that the words *thirty-four*, for example, map onto a number that looks like *XX* and not one that looks like *X*. Children could detect this correlation by hearing multidigit numbers named while also seeing them in print, as they might when parents are commenting on a calendar, asking their child to push the buttons on an elevator, or looking for a room number in an office building. Other statistical cues to number identity include the word order of number names (e.g., *seventy-eight* maps onto *78* better than the name *eighty-seven* in left-right reading order) and the association of single-digit numerals to their verbal names (e.g., the numeral *324* is probably not named *six hundred fifty-one* because it has neither a six, a five, or a one in it).

Identifying a word's referent in a complex perceptual scene is the first step toward determining its meaning, but statistics could also support inferences about multidigit number meanings. For example, children could use their knowledge of single-digit number meanings to make guesses about the ordinality of multidigit numbers. Preschool children know large single-digit numerals (e.g., 7, 8, 9) represent larger quantities than small single-digit numbers (e.g., 1, 2, 3), even if they have not learned the precise meanings of these symbols (LeCorre & Carey, 2007; Sarnecka & Gelman, 2004). This association would be enough to support the correct guess that 899 represents more than 122. Eventually, children may realize that the magnitude of the leftmost digit matters more than the others, or that the number of digits matters more than the meanings of the individual digits. These realizations could stem from experiences with package labels (e.g., a box of blocks labeled XXX is bigger and has more in it than the box labeled XX). Such insights, though not the same as a complete understanding of place value, could be significant steps toward this understanding. However, the possibility that children bootstrap into the place value system has been largely overlooked due to the intense focus on subsequent misconceptions and limitations (e.g., error-prone multidigit calculation).

In fact, there is emerging evidence that young children know quite a lot about multidigit numerals. In a recent study, preschool children demonstrated the ability to write multidigit numerals with some success (Byrge, Smith, & Mix, in press). Moreover, their errors were "intelligent"—appearing to reflect inferences about the structure of written notation gleaned from statistical patterns in verbal input. For example, rather than using spatial position alone to indicate base-10 units in written numerals, children frequently invent a sort of expanded notation, such that *113* is written as *10,013*. Although this error has been viewed by some as further evidence of children's struggles with place value notation (e.g., Bussi, 2011), Byrge et al. (in press) argued that it is more akin to grammatical overgeneralizations (e.g., *goed* instead of *went*) and may represent an important developmental milestone even if it does not align completely with cultural conventions.

In summary, although children may lack a complete understanding of place value until late elementary school and may well struggle with certain misconceptions without instructional support, it is unlikely they enter school a blank slate. Instead, they may bring partial knowledge of the place value system derived from sensitivity to statistical patterns in multidigit numerals and verbal number names. However, this partial knowledge has not been demonstrated empirically and little is known about its developmental course. By taking a broader, developmental perspective, in which multidigit numerals are seen as akin to other forms of language input, this study addresses that gap. Specifically, rather than embedding our assessment in complex tasks, such as multidigit calculation, we used simple tasks that focused directly on the mappings among written numerals, spoken numbers, quantities of dots, and block patterns. Thus, we could more sensitively ascertain the extent to which young children actively make sense of the place value system.

Experiment 1

The first experiment examined the performance of kindergarten, first-, and second-grade children in two tasks: (a) mapping spoken number names to digits, dots, or block representations and (b) indicating which of two arrays—digits, dots, or blocks—represented "more."

Method

Participants

The total sample of 91 children was divided into three age groups based on grade in school (kindergarten: M = 61 months, range = 49–73 months, n = 31; first grade: M = 74 months, range = 61–84 months, n = 25; second grade: M = 86 months, range = 72–96 months, n = 35). Roughly half the participants (n = 37) were boys, and they were evenly divided among the three grades (kindergarten: n = 12; first grade: n = 12; second grade: n = 13). An a priori power analysis, using the statistical program G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009), indicated that a sample size of 21 participants per age group (or n = 63 total) would be adequate to achieve 80% power with a medium effect size.

All children came from the same, ethnically diverse, middle-socioeconomic-status (SES) population, for which the 2010 U.S. Census reported 78.4% European American, 10.6% Asian, 6.8% African American, 1.4% from other ethnicities, and 2.9% from two or more ethnicities. Within these groups, 3.4% were identified as Hispanic or Latino. The median family income in this community was $81,158.

Materials and Procedures

Children's knowledge of multidigit numerals was assessed using two tasks. For the *Which is x?* task, children saw two quantities represented with either (a) written numerals, (b) base-10 blocks, or (c) clouds of dots, and were asked to point to the quantity named by the experimenter (e.g., "Which is one hundred thirty-two?"). The same 20 items were used in each stimulus condition (numerals, blocks, and dots), but they were presented in one of two random orders. The order of the three stimulus conditions was counterbalanced across children.

The numeral displays consisted of two numerals printed in black ink, 96-point Calibri font, and placed side by side in the center of a horizontally oriented 8 × 11 in. sheet of white paper. The two numerals were separated by 2.6–4 in. of white space, depending on how many digits were involved (i.e., one-digit numerals were spaced farther apart than two- and three-digit numerals). Across the 20 trials, the larger of the two quantities was presented on the right for 10 trials and on the left for 10 trials, in a random order.

The blocks displays showed the same quantities, also presented side by side on a horizontal 8 × 11 in. sheet of white paper, but represented using photographs of base-10 blocks. The blocks were lined up left to right, from highest to lowest place, just as they are in written numerals. The overall length of each quantity was 15.65 mm on average (range = 3.81–29.46 mm). As for the numerals condition, the larger quantity appeared on the left for half the trials, and the order of left-right presentation was randomized.

The dot displays contained two arrays of black dots (dot diameter = 1.3 mm), arranged side by side on an 8 × 11 in. sheet of white paper, and separated by a vertical line. Within each display, the dots were randomly dispersed within an imaginary 93 × 70 mm rectangle (total area = 6,510 mm). The sizes of the individual dots varied depending on how many dots were in the display, such that more numerous displays contained smaller individual dots (range = 1.3–3.0 mm). In this regard, number varied inversely with dot size and density, but overall area of the dot clouds was controlled. This was acceptable because our aim was not to test whether children could map or compare exact numbers of discrete dots. Instead, we aimed only to test whether children could (a) map numerals to an approximate quantitative referent and (b) compare such quantities as a baseline for interpreting their performance in the other conditions. Thus, it was not important to control dimensions of continuous amount and, given the set sizes involved, it was preferable to allow access to these dimensions if it helped children discriminate the sets. As for the other stimulus conditions, the larger quantity appeared on the left for half the trials, and the order of left-right presentation was randomized.

The *Which is more?* task was similar except that instead of being asked to identify a specific number (e.g., 58), children were asked to identify which numeral, block configuration, or dots picture represented the larger quantity (e.g., “Which is more?”). The parameters of the displays were the same across tasks, however. For example, the dot displays in the *Which is more?* task contained two arrays of black dots (dot diameter = 1.3 mm, range = 1.3–3.0 mm), arranged side by side on an 8 × 11 in. sheet of white paper and separated by a vertical line. Within each display, the dots were randomly dispersed within an imaginary 93 × 70 mm rectangle (total area = 6,510 mm). As before, the same 20 items were used for all three stimulus conditions (numerals, blocks, and dots) but in different random orders. The order of three stimulus conditions within each task was counterbalanced across children, and the order of the two tasks was counterbalanced such that half the children within each grade completed the *Which is x?* task first.

Because children completed three versions of the 20 *Which is x?* items, and three versions of the 20 *Which is more?* items, they completed 120 items total. Although this constituted a rather lengthy test in terms of number of items, children required only about 10–15 min to complete it and there were no signs of fatigue—that is, all children completed all items. Testing took place in the late fall or early winter, when children had received about 3 months of grade level instruction. For the majority of children, this instruction was based on the *Everyday Mathematics* curriculum. Children were tested in groups of two or three, but were seated in such a way that they were unable to view their peers’ test papers.

The sets of items for the *Which is more?* and *Which is x?* tasks were selected to be roughly comparable but not identical, so that they could be used as within-subject measures in some experiments and because the likely source of errors (and their diagnosis) would benefit from somewhat different comparisons. Both sets included some one-digit numerals so that we could assess whether even young preschoolers could map names to single-digit numbers and compare those numbers with respect to quantity. Furthermore, the *Which is x?* set included numbers that differed only in the addition or place of zero since pilot work indicated one early common error in mapping names to numbers was knowing how to interpret 0. For the *Which is more?* task, pilot work indicated that children generally picked the string with more digits as “more” and therefore the set tested included more (but not all) equal length strings and strings composed of the same digits in different places or that different in just one digit.

Results and Discussion

Children’s proportions correct by task and grade level are presented in Table 1. An inspection of the means suggests there is improvement from kindergarten to second grade, but that all children performed well above chance. This pattern was confirmed using t tests that compared each age group’s mean score to chance (i.e., 50%). Actual performance was significantly different from chance across age, task, and condition (all ps < .001). We examined group differences using a repeated measures analysis of variance (ANOVA) with task (*Which is x?* vs. *Which is more?*) and stimulus condition (numerals, dots, and blocks) as within-subjects factors, and grade level (K, 1, 2) as a between-subjects factor. As expected, there was a significant main effect of grade, $F(2, 264) = 94.37$, mean square error $MSE = .03$, $p < .001$, $\eta_p^2 = .42$, due to significant improvement on the tasks with increasing age such that second graders ($M = 0.87$, $SD = .10$) outperformed both first graders ($M = 0.81$, $SD = .12$), $t(178) = 4.02$, $p < .001$, and kindergarten students ($M = 0.66$, $SD = .15$), $t(196) = 11.77$, $p < .001$, and first graders outperformed kindergarteners, $t(166) = 6.87$, $p < .001$, all two-tailed, Bonferroni t tests.

The ANOVA also revealed a significant main effect of stimulus condition, $F(2, 264) = 34.44$, $MSE = .03$, $p < .001$, $\eta_p^2 = .21$, that was due to better performance in the numerals condition (numerals vs. blocks: $p < .001$; numerals vs. dots: $p < .0001$; blocks vs. dots: $p = .65$) and a significant main effect of task type, $F(1, 264) = 97.51$, $MSE = .01$, $p < .0001$, $\eta_p^2 = .27$, that reflected higher scores overall on the *Which is more?* task (vs. *Which is x?*). These effects were mediated by significant interactions between task and stimulus, $F(2, 264) = 29.73$, $MSE = .01$, $p < .001$, $\eta_p^2 = .18$, and age and stimulus, $F(4, 264) = 3.05$, $MSE = .03$, $p = .02$, $\eta_p^2 = .04$. Pairwise comparisons indicated that the Task × Condition interaction was due to several significant condition differences in the *Which is x?* task ($M_{\text{numerals}} = .86$, $SD = .17$; $M_{\text{blocks}} = .68$, $SD = .68$), numerals versus blocks, $t(180) = 6.95$, $p < .001$, and numerals versus dots, $t(180) = 9.32$, $p < .0001$, Bonferroni two-tailed, but no significant stimulus differences in the *Which is more?* task.

The Age × Stimulus interaction was due to older children (first and second graders) performing

Table 1
Proportion Correct by Task, Stimulus, and Grade (Experiment 1)

Grade	Stimulus condition	*Which is x?*	*t* test	*Which is more?*	*t* test
Kdg	Blocks	.55 (.12)	24.43***	.69 (.21)	23.48***
	Dots	.59 (.09)	35.04***	.71 (.20)	27.00***
	Numerals	.70 (.19)	19.85***	.70 (.24)	20.79***
	Overall	.61 (.15)	7.29***	.70 (.21)	9.20***
First	Blocks	.68 (.14)	23.17***	.86 (.12)	36.47***
	Dots	.65 (.14)	22.82***	.84 (.09)	44.14***
	Numerals	.93 (.09)	47.66***	.89 (.13)	32.55***
	Overall	.75 (.18)	12.32***	.86 (.11)	27.96***
Second	Blocks	.80 (.16)	28.53***	.91 (.09)	61.01***
	Dots	.72 (.12)	34.88***	.88 (.06)	87.37***
	Numerals	.96 (.05)	103.58***	.97 (.04)	115.07***
	Overall	.83 (.16)	21.53***	.92 (.08)	55.19***

Note. One-tailed *t* tests compared each proportion correct to chance level of .50. Kdg = kindergarten.
$^{***}p < .001$.

significantly better in the numerals condition than in both blocks and dots–numerals vs. blocks: first grade $M_{\text{numerals}} = .91$, $SD = .10$; $M_{\text{blocks}} = .77$, $SD = .10$), $t(48) = 4.72$, $p < .001$, and second grade ($M_{\text{numerals}} = .97$, $SD = .04$; $M_{\text{blocks}} = .86$, $SD = .10$), $t(68) = 6.05$, $p < .001$; numerals vs. dots: first grade ($M_{\text{dots}} = .74$, $SD = .08$), $t(48) = 6.32$, $p < .001$, and second grade ($M_{\text{dots}} = .80$, $SD = .06$), $t(68) = 13.66$, $p < .001$, Bonferroni two-tailed–but only a marginally significant condition difference between numerals and blocks for kindergarten students ($M_{\text{numerals}} = .70$, $SD = .19$; $M_{\text{blocks}} = .62$, $SD = .13$), $t(60) = 1.87$, $p < .067$, Bonferroni, two-tailed. Children in second grade also performed better in the blocks condition (vs. dots), $t(68) = 2.90$, $p = .005$, but this difference was not obtained in the other two grades (both ps > .32).

Certain aspects of these results are not surprising in light of previous research. First, there is reason to think children would perform quite well in the *Which is more?* dots and blocks conditions based on studies showing that infants and young children can discriminate large quantities presented in visual arrays (e.g., Cantlon, Platt, & Brannon, 2009; Halberda & Feigenson, 2008; Lipton & Spelke, 2003). Also, we did not prevent children from using continuous perceptual variables, such as density or area, to discriminate these displays so some degree of success was anticipated. Also, it is not surprising that *Which is more?* performance would be higher than the performance in *Which is x?* task given that this task requires a less precise grasp of the quantities being compared. Children only need to identify which quantity is greater. The fact that second-grade students performed better for blocks than dots likely reflects greater familiarity with these materials after exposure to them in school. Most kindergarten and first-grade students would have had far less, if any, experience with base-10 blocks.

What is surprising is that (a) children of all ages performed significantly *better* in the numerals version of both tasks than they did with either of the visual arrays (i.e., blocks or dots) and (b) children of all ages performed above chance on the *Which is x?* version of the blocks and dots conditions even though this task required them to interpret verbal number names. This evidence of strong performance in the conditions and tasks that involve numerals and number names suggests that children understand these symbols well enough to support accurate and precise comparisons.

We next examined children's performance on the numerals tasks more closely, to determine which particular items were relatively accessible. The 20 *Which is x?* items are presented in Table 2, rank ordered for difficulty based on the kindergarten children's mean performance. Recall that there was a 50% probability of being correct by guessing in this forced-choice task. When we compared each age group's performance to chance, item by item, we found that kindergarteners performed randomly on the first 6 items in the table, but performed significantly above chance on the rest.

We can infer the strategies children used by examining the demands and affordances of various items. Several of the easiest items could be solved correctly by recognizing individual written digits (e.g., 2 vs. 8; 4,279 vs. 6,358). That is, if a child knew what the numeral 2 looked like, it would be possible to identify even a four-digit number with 2 in it versus one without. However, other items

Table 2
Proportion Correct by Item and Grade for Which Is x? Numerals (Experiment 1)

	Grade		
Items	Kdg	First	Second
206 vs. 260	0.39	0.96	0.94
356 vs. 536	0.52	0.96	0.94
350 vs. 305	0.58	0.76	0.89
36 vs. 306	0.58	0.92	0.97
2,843 vs. 2,483	0.61	0.80	0.94
267 vs. 627	0.61	0.92	0.97
201 vs. 21	0.65	0.96	0.97
670 vs. 67	0.65	1.00	0.91
85 vs. 850	0.68	0.92	1.00
64 vs. 604	0.68	0.96	0.91
1,002 vs. 1,020	0.68	0.80	0.94
402 vs. 42	0.71	0.88	1.00
1,000 vs. 100	0.74	1.00	0.94
105 vs. 125	0.74	0.96	0.97
11 vs. 24	0.77	1.00	1.00
12 vs. 22	0.77	0.88	1.00
807 vs. 78	0.77	0.92	0.97
4,279 vs. 6,358	0.81	0.92	1.00
15 vs. 5	0.81	1.00	1.00
2 vs. 8	0.84	1.00	1.00

Note. Kdg = kindergarten.

Table 3
Proportion Correct by Item and Grade for Which Is More? Numerals (Experiment 1)

	Grade		
Items	Kdg	First	Second
614 vs. 461	0.55	0.64	0.91
5,687 vs. 8,657	0.55	0.76	0.94
14 vs. 41	0.55	0.84	0.94
16 vs. 62	0.65	0.92	0.94
30 vs. 60	0.65	0.92	0.97
11 vs. 19	0.65	0.96	1.00
123 vs. 321	0.68	0.72	0.97
458 vs. 845	0.68	0.72	0.97
6,892 vs. 2,986	0.68	0.88	0.97
670 vs. 270	0.68	0.96	1.00
26 vs. 73	0.71	0.96	0.97
3 vs. 7	0.71	1.00	1.00
4,620 vs. 4,520	0.74	0.92	0.94
6 vs. 8	0.74	0.96	1.00
101 vs. 99	0.81	1.00	0.94
100 vs. 10	0.81	1.00	0.97
223 vs. 220	0.84	0.92	1.00
72 vs. 27	0.87	0.84	1.00
585 vs. 525	0.87	0.96	0.97

Note. Kdg = kindergarten.

required at least partial knowledge of the place value system. For example, to identify 670 (vs. 67), children could not simply choose the written number that had 6 and 7 in it. They would also have to know that the word *hundred* in the number name signaled more digits. Indeed, kindergarten children performed quite well on a number of items with this requirement (e.g., 850 vs. 85, 402 vs. 42, etc.). Another heuristic children might use in this task would be mapping the first number name to the leftmost digit (as for 807 vs. 78). However, for most items requiring this inference, children performed at chance (e.g., 356 vs. 536) so it is not clear they applied a leftmost digit strategy. Items that required accurate mapping, place by place, and could not be solved using any of the partial knowledge heuristics (e.g., 350 vs. 305; 2,843 vs. 2,483) were among the most difficult for kindergarten students. These items remained the most challenging in first grade, but by second grade, children responded to even these items at ceiling.

The rank ordering of performance on *Which is more?* items is presented in Table 3. On this task, children saw two written numerals and chose the one that represented a larger quantity. Here again, children performed mostly at ceiling by first grade. Kindergarten children performed at chance on the first three items, but were significantly above chance beyond that, albeit not close to ceiling. Thus, at least some kindergarten children had enough knowledge of place value and number meanings to make educated guesses.

Some of the easiest items could be solved by recognizing individual digits (e.g., 6 vs. 8). For a few items, children could answer correctly if they knew that two-digit numbers represent smaller quantities than three-digit numbers (e.g., 101 vs. 99). Others had the same number of digits but one digit differed (e.g., 525 vs. 585; 4,520 vs. 4,620). Although the differences in these pairs was relatively subtle, children performed quite well, perhaps because they tend to choose quantities with more large numbers in them—for example, choosing the number with 8 in it for 525 vs. 585. The more challenging items had the same digits in different orders. It is possible for children with partial knowledge of place value to choose the larger quantity if they understand that the leftmost digit carries more weight. In some cases, they appeared to use this heuristic (e.g., 27 vs. 72, 123 vs. 321, 6,892 vs. 2,986); however, this was not completely consistent as several such items were also among the most difficult (e.g., 614 vs. 461, 5,687 vs. 8,657). Several of this type also remained the most difficult items in first grade.

Overall, the results of Experiment 1 are unexpected given the well-documented difficulties children have grasping multidigit numerals in previous research. This is not to say children have mastered place value at this age, but rather that they know more than previous research might lead one to believe. Even kindergarten students who are unlikely to have received direct instruction in place value can accurately identify and compare written numerals—better, in fact, than they can identify and compare pictorial representations of the same quantities. This indicates that children are actively making sense of multidigit numerals they encounter in everyday life. Direct instruction likely enhances this ability. Children in first and second grades had likely been taught to read and write multidigit numerals (Common Core State Standards Initiative, 2010; National Council of Teachers of Mathematics, 2000) and they reached ceiling on both numerals tasks. Still, children bring a stronger experiential foundation to place value instruction than educators may realize.

The strong performance of kindergarten students in Experiment 1 raises the question of whether even younger children also exhibit understanding of multidigit numerals and large quantities. Perhaps children begin to make certain inferences even earlier. We investigated this possibility in Experiment 2.

Experiment 2

Method

Participants

The total sample of 92 children was divided into three age groups (3½-year-olds: M = 45 months, range = 31–51 months, n = 26; 4½-year-olds: M = 56 months, range = 52–60 months, n = 32, 5-year-olds: M = 65 months, range = 60–72 months, n = 34). Roughly half the participants (n = 44) were boys, and they were evenly divided among the three age groups (3½-year-olds: n = 11; 4½-year-olds: n = 17; 5-year-olds: n = 16). An a priori power analysis, using the statistical program G*Power 3.1 (Faul et al., 2009), indicated that a sample size of 21 participants per age group (or n = 63 total) would be adequate to achieve 80% power with a medium effect size. All children came from an ethnically diverse, middle-SES population similar to that sampled in Experiment 1 but from another state. According to the 2010 U.S. Census, this community was 83.0% European American, 8.0% Asian, 4.6% African American, 1.6% from other ethnicities, and 3.0% from two or more ethnicities. Within these groups, 3.5% reported being of Hispanic or Latino descent. The median family income was $50,054. Participants were recruited from 12 different day cares with diverse programs but with an overall focus on social and play activities. Many of the 5-year-olds were in some form of half-day kindergarten (at their day care); kindergarten is not required by the state of Indiana and the curriculum varies considerably across different schools.

Materials and Procedures

The materials and procedures were the same as in Experiment 1, except that children were tested individually and the tests had fewer items. Specifically, instead of 20 items per test (*Which is x?* vs. *Which is more?*), there were 8 items per test in Experiment 2. Also, children were tested only on the numerals version of both tasks, and not the other conditions (blocks and dots). Thus, children completed 16 items total (vs. 120 in Experiment 1). The number of items was limited because pilot testing indicated that 3-year-olds were unable to complete 120 items. Also, because we were probing the very earliest emergence of competence, we did not include items we knew were more difficult for kindergarten students. Finally, because our main goal was to measure children's ability to interpret written numerals, we tested only numerals items. The 8 items on each test were drawn from a pool of 16 items and counterbalanced across children. To maximize comparability across experiments, the 16-item pool comprised a subset of the 20 items from Experiment 1.

Results and Discussion

Children's proportion correct on the two tasks, by condition and age group, is presented in Table 4. As

Table 4
Proportion Correct by Task and Grade (Experiment 2)

Age in years	*Which is x?* task	t test	*Which is more?* task	t test
3½	.61 (.19)	3.15**	.57 (.14)	2.79**
4½	.69 (.20)	5.37***	.64 (.16)	4.80***
5	.79 (.22)	7.69***	.75 (.15)	9.61***

Note. One-tailed t tests compared to chance level of .50. These data include numerals performance only as blocks and dots were not included in Experiment 2.
$^{**}p < .01$, $^{***}p < .001$.

before, there was significant improvement with age. A repeated measures ANOVA with task (*Which is x?* vs. *Which is more?*) and age (3½, 4½, and 5 years old) as a between-subjects factor revealed a significant main effect of age, $F(1, 89) = 13.59$, $MSE = .04$, $p < .001$, $\eta_p^2 = .23$. Post hoc tests with Bonferroni correction showed that the 5-year-olds performed better than either 3½-year-olds ($p < .001$) or 4½-year-olds ($p = .004$) but 4½-year-olds did not outperform 3½-year-olds ($p = .19$). Still, all three age groups performed significantly above chance (see Table 4). Neither the main effect of task, nor task and age interaction reached significance.

The rank order performance of children on both tasks is presented in Tables 5 and 6. As in Experiment 1, children demonstrated some competence interpreting multidigit numerals from a much earlier age than previous research would suggest. They began to correctly identify and compare two- and three-digit numerals starting at 3½ years of age. Although there was steady improvement from 3½ to 5 years and children were not at ceiling on most items even at age 5, they nonetheless performed significantly above chance on both tasks by 3½ years of age and on most items within each task by 4½ years of age. This is most likely attributable to exposure to multidigit numerals without formal instruction as not all children in this study attended preschool and even for those who did, place value instruction probably was not offered. Thus, as in other areas of language development, it appears children infer the meanings of these numbers using whatever experiences they can access. If we focus on only the *Which is x?* items for which 3½-year-olds performed above chance, it appears they used two strategies. One strategy was simply recognizing familiar written digits, like 2, 8, and 12. The other strategy was knowing that the word *hundred* in a number name signals more digits (e.g., 201 vs. 21). This was not consistent, however, and most items that could be answered correctly that way were missed by 3½-year-olds (e.g., 670 vs. 67). Among 4½-year-olds, this pattern was more consistent albeit far from ceiling.

Indeed, looking across Experiments 1 and 2, children did not reach ceiling until second grade, by which time they had almost certainly received place value instruction in school. Now that we know more specifically where younger children's difficulties lie (i.e., which particular items were most difficult for them as evident in Tables 2, 3, 5, and 6), it is interesting to ask whether place value instruction can improve learning on these items. Previous research suggests two main approaches to place value instruction. One focuses on written symbols (Kamii, 1986) whereas the other incorporates concrete models, such as base-10 blocks (Fuson & Briars, 1990). In Experiment 3, we provided both types of instruction to a group of kindergarten students to see whether either approach would lead to improvement.

Table 5
Proportion Correct by Item and Age in Years for Which is x? Numerals (Experiment 2)

	Age in years		
Items	3½	4½	5
206 vs. 260	0.42	0.50	0.59
1,000 vs. 100	0.46	0.38	0.68
807 vs. 78	0.50	0.72	0.74
670 vs. 67	0.42	0.69	0.76
1,002 vs. 1,020	0.62	0.75	0.76
201 vs. 21	0.73	0.66	0.79
64 vs. 604	0.69	0.69	0.79
105 vs. 125	0.46	0.75	0.79
350 vs. 305	0.54	0.75	0.79
85 vs. 850	0.50	0.78	0.79
402 vs. 42	0.54	0.66	0.85
36 vs. 306	0.54	0.69	0.85
15 vs. 5	0.58	0.97	0.88
11 vs. 24	0.62	0.81	0.88
12 vs. 22	0.73	1.00	0.91
2 vs. 8	0.73	1.00	0.91

Table 6
Proportion Correct by Item and Age in Years for Which is More? Numerals (Experiment 2)

	Age in years		
Items	3½	4½	5
670 vs. 270	0.50	0.47	0.56
4,620 vs. 4,520	0.50	0.53	0.65
26 vs. 73	0.62	0.56	0.65
14 vs. 41	0.62	0.59	0.68
72 vs. 27	0.42	0.47	0.71
123 vs. 321	0.46	0.59	0.71
223 vs. 220	0.69	0.72	0.71
101 vs. 99	0.62	0.53	0.74
16 vs. 62	0.42	0.66	0.74
585 vs. 525	0.46	0.63	0.76
30 vs. 60	0.58	0.69	0.76
6 vs. 8	0.58	0.81	0.85
11 vs. 19	0.73	0.66	0.91
3 vs. 7	0.69	0.91	0.94
100 vs. 10	0.73	0.72	0.94

Experiment 3

Method

Participants

A sample of 24 kindergarten students (M = 57 months, range = 45–69 months) was recruited from the same, ethnically diverse, middle-SES population sampled in Experiment 1. Roughly half the participants (n = 13) were boys. Children were evenly divided into two training groups: base-10 blocks (n = 12) and symbols-only (n = 12). An a priori power analysis, using the statistical program G*Power 3.1 (Faul et al., 2009), indicated that a sample size of 21 participants would be adequate to achieve 80% power with a medium effect size.

Materials and Procedure

Children's knowledge of multidigit numerals was assessed using the two tasks from Experiment 1 (i.e., *Which is x?* and *Which is more?*). All three conditions (numerals, dots, and blocks) were tested within each task. There were 20 items from each condition, resulting in 60 items per task or 120 items total. Children completed both tasks prior to training and then again within 2 days of the final training session.

The two training groups completed five lessons on topics that included sorting and matching block types or number cards, copying block patterns or numerals, mapping numerals to blocks, mapping blocks to numerals, and block and number comparisons. The content and instructional approaches were modeled on those commonly used in schools and typical mathematics curricula (e.g., *Everyday Mathematics*). The two training conditions were based on different theoretical notions about the emergence of symbolic meaning and symbol grounding. One idea is that symbols are grounded in concrete experience (Barsalou, 2008; Lakoff & Nunez, 2000) and these experiences may be especially critical for young children (e.g., Bruner, Olver, & Greenfield, 1966; Cuisenaire & Gattegno, 1953; Dienes, 1961). In support of this, studies have shown improved mathematical performance for children taught using concrete models or manipulatives (e.g., Fuson & Briars, 1990; see Sowell, 1989, for a review). An alternative view is that children construct an understanding of complex symbols via experiences with symbols themselves via logico-mathematical processes (e.g., Chandler & Kamii, 2009; Kamii, 1986). Even theories of grounded cognition admit a role for chaining from one symbolic representation to another (e.g., Lakoff & Nunez, 2000). That is, once the building blocks of a symbol system have been grounded in direct experience (e.g., the meaning of single digits), more complex symbolic forms can be understood with reference to these building blocks, rather than grounding every instance in a concrete experience. In line with this view, some researchers have claimed concrete models are either ineffective or detrimental (Ball, 1992; Kaminski, Sloutsky, & Heckler, 2008; McNeil, Uttal, Jarvin, & Sternberg, 2009). If early knowledge is based on intuitive awareness of the statistical regularities across written and spoken number names, then training that amplifies the regularities in these symbol systems could be the best approach to improving explicit understanding of place value. Alternatively, training that connects children's early and not completely correct knowledge of place value to its conceptual underpinnings might be more effective, and such training could benefit from bridging via concrete representations, such as base-10 blocks. Accordingly, our training compared both kinds of approaches.

The lessons were presented by a highly trained experimenter over a period of 2 weeks, for 225 min of instruction in total. Children worked in groups of four students but each had his or her own set of materials. The lesson content was identical for both groups except that children in the base-10 blocks group worked out the problems using individual sets of base-10 blocks (15-ones, 15-tens, 15-hundreds, and 2-thousands blocks). They also received vinyl mats (World Class Learning Materials, Chandler, NC) that were divided into four sections by place (thousands, hundreds, tens, and ones). The sections were arranged horizontally, from thousands on the left to ones on the right, thereby mirroring the order of written multidigit numerals. Also, each section was illustrated with a line drawing of the corresponding base-10 block. Children in the symbols-only condition were given plain white note cards with a single hand-written numeral ranging from zero to nine to complete parallel activities. For example, if children in the base-10 blocks group were asked to represent a multidigit numeral using blocks, children in the symbols-only group did so using their digit cards.

Results and Discussion

Children's combined pretest scores were equivalent across conditions ($M_{\text{symbols-only}}$ = .68, M_{blocks} = .69), $t(70)$ = .03, p = .98, two-tailed, so children entered the study at the same level of place value knowledge. At pretest, children in both groups performed significantly above chance on the numerals items, but not on blocks or dots, on the *Which is x?* task and the opposite was true for the *Which is more?* task. (see Table 7). Both groups performed significantly above chance on average following training: symbols-only (M = 0.62, SD = .10), t(35) = 7.29, p < .001, and blocks (M = 0.54, SD = .10), t(35) = 2.36, p = .02, and within some tasks and conditions, (see Table 7). However, pairwise comparisons indicated that the gains from pre- to posttest were not significant for any task in either condition (all ps > .05). Indeed, for blocks children, scores in the numerals condition actually dropped from pretest (M = 0.58, SD = .07) to posttest (M = 0.50, SD = .06), t(11) = 4.28, p = .001.

To further examine the differential effects of training, we submitted children's posttest scores to a repeated measures ANOVA with task (*Which is x?* vs. *Which is more?*) and condition (numerals, dots, and blocks) as within-subject variables and training condition (blocks vs. symbols-only) as the between-subjects factor. There was a significant main effect of task, F(1, 70) = 212.37, p < .001, η_p^2 = .75, such that *Which is more?* posttest scores were higher ($M_{\text{which-is-x?}}$ = .51, SD = .13; $M_{\text{which-is-more?}}$ = .57, SD = .09), t(23) = 1.80, p = .04, one-tailed; however, there were no other main effects or interactions. Thus, it appeared that neither training condition led to improved scores.

If we focus on only the items that were particularly challenging, a different pattern emerged. We carried out a second ANOVA focusing on the numerals items for which children originally performed at chance because we wanted to determine whether training addressed the particular limitations children exhibited on school entry. The resulting 27 items (13 *Which is more?* and 14 *Which is x?*) were submitted to a repeated measures ANOVA with test type (pre vs. post) as the within-subject variable and training condition (blocks vs. symbols-only) as the between-subjects variable. First, there was a significant main effect of test, F(1, 22) = 7.62, MSE = .008, p = .01, η_p^2 = .26, that reflected overall improvement after training (M_{pretest} = .49, SD = .11; M_{posttest} = .56, SD = .12), t(23) = 2.55, p = .009, Bonferroni, one-tailed. But there also was a Test × Condition interaction, F(1, 22) = 5.03, MSE = .04, p = .01, η_p^2 = .19, in favor of the symbols-only training group (M_{pretest} = .48, SD = .14; M_{posttest} = .61, SD = .09). In short, there was a significant increase from pre- to posttest for the symbols-only group, t(11) = 4.24, p = .001, but not for those who received training with blocks (M_{pretest} = .50, SD = .10; M_{posttest} = .51, SD = .13), t(11) = .32, p =

Table 7
Proportion Correct by Task and Training Condition (Experiment 3)

	Which is x?						*Which is more?*					
	Numerals		Blocks		Dots		Numerals		Blocks		Dots	
Condition	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Symbols-only	0.63 (.19)	0.72 (.11)	0.53 (.08)	0.54 (.11)	0.53 (.10)	0.57 (.09)	0.53 (.12)	0.55 (.12)	0.68 (.11)	0.64 (.17)	0.66 (.12)	0.7 (.16)
t test	2.5*	7.10*	1.07	1.19	0.88	2.68*	0.83	1.55	5.6*	2.84*	4.63*	4.13*
Blocks	0.60 (.08)	0.54 (.13)	0.52 (.07)	0.55 (.15)	0.53 (.13)	0.50 (.10)	0.55 (.08)	0.45 (.13)	0.62 (.11)	0.60 (.18)	0.67 (.10)	0.58 (.12)
t test	4.5*	1.03	0.89	1.17	0.80	0.14	1.89	1.23	3.68*	1.84	5.73*	2.50*

Note. One-tailed *t* tests compared each proportion correct to chance level of .50.
*p < .05.

.75. Thus, there was evidence that symbols-only training led to improvement on certain items, but no evidence that training with base-10 blocks was helpful.

One could argue that these training effects are due to improvement that would have occurred over the period of the study even without training. It seems unlikely that the degree of improvement we observed could naturally occur over a 3-week period, given that it took children roughly 3 years to progress from emergent competence to mastery in Experiment 1. Also, if the results were due to maturation or incidental learning, we should not obtain different patterns of improvement for the two training conditions. Still, without a no-training comparison group, it is impossible to rule out this interpretation entirely.

General Discussion

Both research and the observations of teachers indicate that place-value notation is difficult for school age children to learn. Yet the present results indicate that preschool children know enough about how large numbers are written to map them to their spoken names and to judge relative magnitudes. To do this, they must know—at the very least—that place matters: The same digit represents more when it is to the left then to when it to the right in a string of digits (e.g., Which is more: 123 vs. 321?), numbers are read left to right (e.g., Which is two-hundred-sixty-seven? 267 vs. 627?), and the number of digits is related to both magnitude (e.g., Which is more: 101 vs. 99?) and identity (e.g., Which is four-hundred-two: 402 vs. 42?). Children's ability to distinguish pairs such as 64 versus 604 and 21 versus 201 also suggests they know something about zeros as place holders. This is certainly a much more extensive conceptual foundation than previously believed.

Because most children in the study had not received formal schooling with multidigit numbers, it seems likely they developed this knowledge from being in a literate world with written numbers. Their encounters with written multidigit numerals, coupled with knowledge of single-digit number meanings, exposure to spoken multidigit number names, and experience with a rough ordering of multidigit numerals (e.g., as one could observe in street addresses), could reveal several statistically regular patterns in this representational system—regularities children could use to make inferences and educated guesses about the meanings of multidigit numerals. We know children glean such regularities from the broad stream of language input and use these regularities to discern word boundaries (Saffran, Aslin, & Newport, 1996), word meaning (Smith & Yu, 2008), grammatical categories (Mintz, Newport, & Bever, 2002), and letter–sound associations (Treiman & Kessler, 2006). For example, Mintz et al. (2002) discovered that the co-occurrence patterns of words in speech to children supplied enough regularity to support discrimination of nouns and verbs. As these studies demonstrate children are sensitive to regularities in written words and speech, we should not be surprised if they also are aware of the regularities in written numerals.

This early competence, however, is far from a deep understanding of multidigit number meaning. There is clearly still much for children to learn about place value in school, and concrete models may play a role in this learning. However, the present results indicate that the potential correspondence between the different sized blocks and place value as realized in spoken number names is not obvious to young preschool children, because they performed relatively poorly in the blocks conditions. Moreover, children consistently performed best in the numerals conditions and tasks that required numeral-to-array mapping. This suggests the intriguing possibility that understanding physical place value blocks requires some prior (albeit incomplete) understanding of the written place value system. Manipulatives thus might not be the entry point for instruction, but better used to augment early understanding, perhaps as a way of making the knowledge latent in the writing system explicit. In contrast, the symbolic representations to which children have been exposed—and about which they have acquired some knowledge—may be the better starting point for explicit instruction.

Consistent with this idea, in Experiment 3, we found that instruction with written symbols led to gains, but not so with base-10 blocks. This adds to the evidence that base-10 blocks are not particularly transparent to children and also lends support to the view that experience with symbols themselves is an important, if not sole inroad to understanding place value (Kamii, 1986). Moreover, whereas blocks training did lead to improvement on blocks outcome measures, it is remarkable that symbols-only instruction also led to significant improvement on these tasks. This suggests that symbols may be an inroad for understanding base-10 blocks as much or more than base-10 blocks are an inroad to understanding written symbols.

Vygotsky argued that conceptual development is supported and transformed by the internalization of cultural tools, such as language. Experiments 1 and 2 suggest that this process begins early, before formal training as preschool children appear to have developed an understanding of the symbols for multidigit numerals via everyday experiences with the symbols themselves. These experiences with written numbers may also affect their understanding of meanings of large numbers and their ability to interpret other representational systems such as base-10 blocks and dots, as measured in the *Which is x?* task. Although this notion may seem antithetical to theories of embodied cognition, which hold that symbols are understood via grounding in perceptual experience, cultural tools have factored prominently in these theories as well (Clark, 1997; Lakoff & Nunez, 2000). For example, Clark (1997) argued that cultural tools, such as numerals, provide scaffolding that offloads task demands and preserves cognitive resources (e.g., attention, working memory, etc.), thereby allowing learners to function at a higher level and, perhaps, achieve new insights. On this view, one could say young children use partial knowledge of multidigit numerals to discover the meaning of large quantities and, thus, bootstrap their way into competence with this symbol system. Experiences with concrete models may depend on this prior knowledge but also feedback on and support a deeper understanding of the conceptual underpinnings of place value. We see this as a critical question for future research.

Regarding subsequent development of place value, one problem with the existing evidence is that place value is often confounded with other concepts and skills. For example, children's accuracy in multidigit calculation undergoes major improvement from 7 to 11 years of age. The fact that children have difficulty with carrying, borrowing, zeroes, and so forth has been taken as evidence that they lack place value concepts (Jesson, 1983). This may be true because these operations require place value knowledge, but they also require knowledge of algorithms themselves, careful execution of these algorithms, accurate retrieval of number facts, and much more. Thus, when children fail to solve multidigit calculation problems, it is not correct to conclude that they lack place value concepts as there are many potential sources of this failure. Indeed, research on addition difficulty has demonstrated that problem size (i.e., magnitude of numbers to be added) accounts for more variance than whether or not carrying is required. (Klein et al., 2010). Similarly, children's inability to represent quantities using base-10 blocks could have more to do with their inability to understand these blocks than it does with their understanding of the way written numerals represent place value.

This study tells us only that young children are developing knowledge about multidigit numbers. It does not provide fine-grained information about the precise nature of that knowledge and its limitations. It also does not identify the origins of this knowledge and the kinds of experiences on which it depends. We have speculated about the streams of input that may be most informative, but these hypotheses are important topics for future research. In addition to providing insight into the development of symbolic ability, this study also informs educational practice. The starting point for the design of any instructional system requires knowing not just the outcome desired but also what children already know, what they will bring—right or wrong—to the instructional experience. The present experiments show young children, prior to formal instruction, have ideas about both spoken number names and written multidigit numbers. The findings point to a critical need to study this knowledge in greater depth, to understand its strengths and weaknesses, and to determine the best way for formal instruction to make contact with and advance that knowledge.

References

Aslin, R. N., & Newport, E. L. (2012). Statistical learning from acquiring specific items to forming general rules. *Current Directions in Psychological Science*, *21*, 170–176. doi:10.1177/0963721412436806

Ball, D. (1992). Magical hopes: Manipulatives and the reform of math education. *American Educator*, *16*, 14–18.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth*. New York, NY: Wiley.

Bussi, M. G. B. (2011). Artefacts and utilization schemes in mathematics teacher education: Place value in early childhood education. *Journal of Mathematics Teacher Education*, *14*, 93–112. doi:10.1007/s10857-011-9171-2

Byrge, L., Smith, L. B., & Mix, K. S. (in press). Beginnings of place value: How preschoolers write three-digit numbers. *Child Development*. doi:10.1111/cdev.12162

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*, 83–91. doi:10.1016/j.tics.2008.11.007

Chandler, C. C., & Kamii, C. (2009). Giving change when payment is made with a dime: The difficulty of tens and ones. *Journal for Research in Mathematics Education*, *40*, 97–118.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. Retrieved from http://www.corestandards.org/the-standards/mathematics

Cuisenaire, G., & Gattegno, C. (1953). *Numbers in color: A new method of teaching arithmetic in primary schools* (2nd ed.). London, UK: Heinneman.

Dienes, Z. P. (1961). *Building up mathematics*. London, UK: Hutchinson Educational.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Fosnot, C. T., & Dolk, M. (2001). *Young mathematicians at work: Constructing number sense, addition and subtraction*. Portsmouth, NH: Heinemann.

Fuson, K. C. (1990). Issues in place-value and multidigit addition and subtraction learning and teaching. *Journal for Research in Mathematics Education*, *21*, 273–280.

Fuson, K. C., & Briars, D. J. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education*, *21*, 180–206.

Graf-Estes, K. M., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*, 254–260. doi:10.1111/j.1467-9280.2007.01885.x

Gunderson, E. A., & Levine, S. C. (2011). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science*, *14*, 1021–1032. doi:10.1111/j.1467-7687.2011.01050.x

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5- and 6-year-olds and adults. *Developmental Psychology*, *44*, 1457–1465. doi:10.1037/a0012682

Ho, C. S. H., & Cheng, F. S. F. (1997). Training in place-value concepts improves children's addition skills. *Contemporary Educational Psychology*, 22, 495–506.

Jesson, D. F. S. J. (1983). The development of place value skills in primary and middle school children. *Research in Education*, *29*, 69–79.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337–374. doi:10.1016/S0010-0285(02)00500-5

Kamii, C. (1986). Place value: An explanation of its difficulty and educational implications for the primary grades. *Journal of Research in Childhood Education*, *1*, 75–86. doi:10.1080/02568548609594909

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, *230*, 454–455. doi:10.1126/science.1154659

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*, 171–184. doi:10.1037/a0025405

Klein, E., Moeller, K., Dressel, K., Domahs, F., Wood, G., Willmes, K., & Nuerk, H. C. (2010). To carry or not to carry—Is this the question? Disentangling the carry effect in multi-digit addition. *Acta Psychologica*, *135*, 67–76. doi:10.1016/j.actpsy.2010.06.002

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology*, *42*, 59–68. doi:10.1037/a0019671

Kouba, V. L., Brown, C. A., Carpenter, T. P., Lindquist, M. M., Silver, E. A., & Swafford, J. O. (1988). Results of the fourth NAEP assessment of mathematics: Number, operations, and word problems. *The Arithmetic Teacher*, *35*, 14–19.

Labinowicz, E. (1985). *Learning from children: New beginnings for teaching numerical thinking*. Menlo Park, CA: Addison-Wesley.

Lakoff, G., & Nunez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

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. doi:10.1016/j.cognition.2006.10.005

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*, 396–401. doi:10.1111/1467-9280.01453

MacNamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review*, *79*, 1–13.

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learning and Instruction*, *19*, 171–184. doi:10.1016/j.learninstruc.2008.03.005

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*, 393–424.

Miura, I. T. (1987). Mathematics achievement as a function of language. *Journal of Educational Psychology*, *79*, 79–82.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood*. New York, NY: Oxford University Press.

Moeller, K., Martignon, L., Wessolowski, S., Engel, J., & Nuerk, H. C. (2011). Effects of finger counting on numerical development—The opposing views of neurocognition and mathematics education. *Frontiers in Psychology*, *2*, 328. doi:10.3389/fpsyg.2011.00328

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Ross, S. H. (1990). Children's acquisition of place-value numeration concepts: The roles of cognitive development and instruction. *Focus on Learning Problems in Mathematics*, *12*, 1–17.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, *92*, 329–352. doi:10.1016/j.cognition.2003.10.001

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. doi:10.1016/j.cognition.2007.06.010

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, *20*, 498–505.

Towse, J. N., & Saxton, M. (1997). Linguistic influences on children's number concepts: Methodological and theoretical considerations. *Journal of Experimental Child Psychology*, *66*, 362–375. doi:10.1006/jecp.1997.2389

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, *98*, 642–652. doi:10.1037/0022-0663.98.3.642

Yu, C., Ballard, D. H., & Aslin, R. N. (2005). The role of embodied intention in early lexical acquisition. *Cognitive Science*, *29*, 961–1005. doi:10.1207/s15516709cog0000_40