DOI: 10.1111/cdev.13724

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

The development of place value concepts: Approximation before principles

Kelly S. Mix[1] | **Corinne A. Bower**[1] | **Gregory R. Hancock**[1] | **Lei Yuan**[2] | **Linda B. Smith**[3]

[1]University of Maryland, College Park, Maryland, USA
[2]University of Colorado, Boulder, Colorado, USA
[3]Indiana University Bloomington, Bloomington, Indiana, USA

Correspondence
Kelly S. Mix, Department of Human Development and Quantitative Methodology, University of Maryland, 3304 Benjamin Building, College Park, MD 20742, USA.
Email: kmix@umd.edu

Funding information
National Science Foundation, Grant/Award Number: DRK-12 1664781 and DRK-12 1621093

Abstract

Place value concepts were measured longitudinally from kindergarten (2017) to first grade (2018) in a diverse sample (n = 279; M_{age} = 5.76 years, SD = 0.55; 135 females; 41% Black, 38% White, 8% Asian, 12% Latino). Children completed three syntactic tasks that required an explicit understanding of base-10 symbols and three approximate tasks that could be completed without this explicit understanding. Approximate performance was significantly better in both age groups. A factor analysis confirmed that syntactic and approximate tasks tapped separate latent variables in kindergarten, but not in first grade. Path analyses indicated that only kindergarten approximate performance predicted overall first-grade place value understanding. These findings suggest that explicit understanding of base-10 principles develops from implicit, partial knowledge of multidigit numbers.

Compositional symbol systems consist of a finite set of symbols and a finite set of rules for combining symbols, also known as syntax, that work together to yield a potentially infinite set of represented meanings. Natural language is one such symbol system, and it is widely known that people can intuitively learn the syntax of natural language without explicit knowledge of the units, combinatorial rules and transformations without formal training. Mathematics is another compositional symbol system, for which children also often begin learning with partial, approximate or intuitive understandings (e.g., Newcombe et al., 2015; Odic et al., 2016). However, unlike natural language, success in mathematics requires explicit knowledge of its representational syntax to perform advanced operations. A critical question is whether, and if so how, this early intuitive understanding of mathematical syntax might support later acquisition of explicit syntactic knowledge.

In the present study, we consider children's informal knowledge and formal learning about the symbol system for representing multidigit numbers. By one hypothesis, early approximate understandings of multidigit numbers could provide a foundation for later explicit understanding of the underlying syntax. By an alternative hypothesis, early approximate knowledge of multidigit numbers could be unrelated or even a hindrance to later learning. In this paper, we specifically examine whether early intuitive knowledge in kindergarten predicts successful initial learning about the representational system for multidigit numbers in first grade. Answering this question is more difficult than it might seem, however, because there are many ways to assess children's understanding of multidigit number meanings and these assessments may not all tap the same underlying understanding.

Abbreviations: ANCOVA, analysis of covariance; CFA, confirmatory factor analysis; CFI, comparative fit index; RMSEA, root mean square error of approximate; WISC–V, Wechsler Intelligence Scale for Children, 5th edition.

What does it mean to understand multidigit numbers?

Multidigit number names express quantities in terms of base-10 units (ones, tens, hundreds, etc.) and counts of these units. To interpret number names and written numerals, and eventually perform multidigit calculations, one needs to know how the units are *related* to each other and the equivalence transformations that those relations support, what we call *count+unit syntax*. We introduce this term to distinguish it from "place value knowledge," which is a general term that could encompass many understandings, and from "base-10," which describes the numeration system we use (i.e., base-10 vs. base-2 or base-5). In our use of the term, *count+unit syntax* refers specifically to the symbolic relations people must unpack to understand what written and spoken number names mean. The first step in that understanding, and the focus of the present study, is how the symbols denote counts of units. Unit counts are reflected in number names, such "three hundred," which represents a count of three for units of 100, or the number name "five thousand," which represents a count of five for units of 1000. Unit counts are also reflected in written numerals. That is, units are denoted by the spatial position in a multidigit numeral (or place), such that ones are to the right, hundreds are just left of ones, thousands are just left of hundreds, and so forth, and the counts of each unit are denoted by the numeral in each place. Thus, the written numeral "532" stands for a count of five units of 100, three units of 10, and two units of 1 (e.g., $532 = [5 \times 100] + [3 \times 10] + [2 \times 1]$). It is possible to deconstruct the syntax of a multidigit number into units and counts, much like one can deconstruct a sentence into nouns, verbs, and other parts of speech. A later step, and not the current focus, is understanding the multiplicative relations among the counted units and the transformations they license. By preschool age, most children know how to count, but in the case of multidigit numbers, the challenge is likely knowing what is being counted (i.e., base-10 units).

Considerable evidence suggests that this first step in learning count+unit syntax is notoriously difficult for school-aged children to master and misunderstandings are common (Fuson, 1988; Mann et al., 2012; Ross & Sunflower, 1995). However, the combinatorial structure—counts of units—form the essence of rules of equivalence transformations. Thus, children who are unable to decode count+unit representational pattern are likely to experience poor performance in mathematics that can persist well into the upper elementary grades (e.g., Carpenter et al., 1997; Raghubar et al., 2009).

Other evidence, however, indicates that many preschool children—some as young as 3 years of age—have informally acquired ideas about multidigit numbers that make them at least *appear* to know something about count+unit syntax (Yuan, Prather, et al., 2019). Many children map multidigit number names to written digits at above chance levels, correctly judging that "five hundred thirty-two" maps onto "532" and not "325" (Mix et al., 2014). They also, at above chance levels, correctly judge relative magnitudes just given the written digits, knowing the "532" is more than "486" (Cheung & Ansari, 2021; Mix et al., 2014). These skills improve across the 3- to 6-year age range, extending to increasingly difficult comparisons and mappings, and larger multidigit numbers. Interestingly, preschool children also fare better on these symbolic comparisons than they do making the same judgments of discrete objects, suggesting that children abstract these relational structures from the symbols themselves (Yuan, Prather, et al., 2019). Preschool children also maintain the order (but not the interval distance) of multidigit numbers up to 1000 reasonably well on the number line estimation task (Yuan, Prather, et al., 2020). Many children can write multidigit numerals from dictation by 6 years of age, and of those who fail to write the conventional form, many produce overregularized errors that reveal perhaps beginning knowledge of the underlying syntax, starting at 4 years of age. For example, when asked to write the numeral for "six hundred and forty-two," some children write "600402" (Byrge et al., 2014; Power & Dal Martello, 1990; Seron & Fayol, 1994). As these authors have argued, the partial understandings observed in preschool are not equivalent to knowing count+unit syntax, but it is striking to see how many of the statistical regularities in base-10 symbols young children have picked up informally and can exploit. An important question is whether these early approximations contribute to children's mastery of count+unit principles when they are formally introduced to them in school.

The problem is there are many ways to measure understanding of multidigit numbers, and it is possible to perform well on some tasks, such as mapping number names to written forms, magnitude comparison, and number line estimation, without having an explicit understanding of the count+unit bases of the symbolic representations. Children could know, for example, that "five hundred and thirty-two" names the numeral "532" and not "325" solely because they know if the word "five" is uttered first, there is likely a "5" in the leftmost position. Likewise, children may infer that "532" signifies more than "53" because they have deduced that 3-digit numerals represent larger quantities than 2-digit numerals, as Cheung and Ansari (2021) have demonstrated. Performance based on approximations such as these are very different from deconstructing multidigit numbers into units and counts. In support of this notion, analyses of children's errors in tasks that permit approximate responding also suggest that their seeming knowledge might be less than it appears (Yuan, Xiang, et al., 2020). For example, many preschoolers can point out the number "100" when the alternative choice is "101," but not when the alternative choice is "1000." If children had explicit understanding of count+unit syntax, it seems likely they would recognize "one hundred" across multiple comparisons. Examinations of item difficulty have also shown that young children fare better on items that allow approximate responses (Cheung & Ansari, 2021; Mix et al., 2014). For example, kindergarteners perform better when asked to compare two numerals with a different number of digits (e.g., 402 vs. 42) than they do when comparing two numerals with the same number of digits (e.g., 402 vs. 316).

An important developmental question is whether the accumulating partial knowledge that is demonstrated in these preschool studies is a conceptual detour that actually distracts children from achieving explicit understanding of count+unit syntax, or a critical stepping stone that contributes to it. There is currently evidence to support both hypotheses. On one hand, as noted above, the evidence indicates that children acquire count+unit syntax slowly and with much difficulty. We know that this difficulty impedes learning because children who lack explicit understanding of count+unit syntax are at greatest risk for low mathematics achievement as they progress through the elementary grades (Chan et al., 2014; Moeller et al., 2011; Ross, 1986). For example, children who make positional errors on a multi-digit number dictation task in first grade tend to earn worse mathematics grades in third grade, as well as exhibiting more frequent place value errors in regrouping problems (Moeller et al., 2011). Positional errors reflect a poor understanding of the representational components, so the fact that children who lack this understanding go on to worse mathematics performance suggests that whatever approximate knowledge they have is not adequate to support later mathematics achievement on its own and may be limited in what it can contribute. Moreover, it stands to reason that if children have shortcuts that allow them to solve most tasks involving multidigit numbers, they would use these less effortful strategies rather than put forth the extra effort needed to decode the syntax of multidigit numbers, and if they are particularly skilled at these shortcuts, may prefer using them for some time. In the same sense that Piaget noted children prefer to assimilate new knowledge to their existing schemas (Piaget, 1952), children who are skilled at approximating multidigit number meanings may also tend to persist, thereby delaying the acquisition of more explicit understandings. In this case, eventual mastery of count+unit syntax may be predicted better by emergent syntactic knowledge than it is by early approximate knowledge, even if better approximate knowledge supports better overall base-10 performance at this age.

On the other hand, approximate knowledge of count+unit syntax, though limited, may nonetheless benefit performance in tasks that require explicit knowledge. Approximate tasks have been widely used in studies showing that early number skills of children in preschool, kindergarten, first grade, and second grade predict later mathematics achievement (Bodovski & Farkas, 2007; Bugden & Ansari, 2011; Duncan et al., 2007; Gersten et al., 2005; Holloway & Ansari, 2009; Jordan et al., 2009; Krajewski & Schneider, 2009a, 2009b; Mazzocco & Thompson, 2005; Mazzocco et al., 2003; Morgan et al., 2009). For example, first graders' arithmetic skills predicted their later mathematical achievement on a standardized assessment in fourth grade (Krajewski & Schneider, 2009a). These findings may reflect one of three things. One possibility is that, unlike preschoolers, children in this age range have acquired an explicit understanding of count+unit syntax that is reflected in their performance on these tasks, even though the tasks do not require it. A second possibility is that children in this age range are mixed in their understanding of base-10 symbols, with some doing well because they have explicit knowledge of count+unit syntax and others using approximate strategies. A third possibility is that performance on these tasks reflects something other than the explicit understanding of count+unit syntax, and it is those approximate understandings themselves that contribute to later mathematics achievement. In order to disentangle these possibilities, an important next step is to assess the same children at the same ages using measures that either permit approximate responding or instead, require explicit knowledge of count+unit syntax. If these measures tap different latent constructs, this finding would indicate there are distinct approaches to base-10 tasks at this age.

The measurement problem

Measures of children's understanding of numbers tend to be significantly inter-correlated suggesting they tap the same underlying construct (e.g., Mix et al., 2016); however, in the case of base-10 symbols, not all measures appear on their face to require the same understandings. For example, when deciding which of two written numerals represents a larger quantity, it likely helps to understand count+unit syntax, but one can also generate correct answers based on a rough understanding of large number meanings. Tasks such as these are the same ones preschool children appear to solve via piecemeal knowledge and thus may not be good indicators of the emerging understanding of count+unit syntax. In contrast, other tasks directly query count+unit syntax, such as telling which digit in a multidigit numeral represents the tens place (Kamii, 1986). Tasks such as these are the same ones elementary students tend to fail and also are highly predictive of later mathematics achievement. To make progress, we selected on a priori grounds, three tasks that by hypothesis can be solved with *approximate* knowledge and three tasks that require knowledge of the compositional components—that is, counts and base-10 units—that are the foundation for *syntactic* knowledge, with the goal of examining their factor structure and underlying relations.

The selected approximate measures are all ones that preschool children, before any formal introduction to base-10 symbols, show some competence: magnitude comparison, transcoding, and number line estimation. As noted above, magnitude comparison ("Which is more, 532 or 356?") may be solved by approximate knowledge; knowing only, for example, the relative magnitudes of the digits 1–9 and additionally that the leftmost digit gets more attention when determining magnitude. Transcoding measures the mapping of number names to written forms through reading and writing and also can be solved through partial knowledge, such as linking the temporal order of spoken words to the left-to-right order of written symbols (Byrge et al., 2014). In the number line estimation tasks (e.g., Siegler & Opfer, 2003; Yuan, Prather et al., 2020), children are asked to place various numerals on a number line that is anchored different ways (0–100, 0–1000, etc.), but children can place multidigit numbers in roughly correct positions without knowing that digits in the decades place stand for a certain number of tens (Krajewski & Schneider, 2009b). Note that by our hypothesis, the underlying processes that permit good performance on these approximate tasks may be a combination of various heuristics and rough approximations, and our goal is not to distinguish among these. Rather, we group these tasks according to what they do *not* require—namely, an explicit understanding of count+unit syntax.

In contrast, the selected syntactic tasks directly probe children's understanding of units and counts. This selection criterion does prevent children with partial knowledge from potentially devising strategies to respond correctly without actually knowing count+unit syntax; however, the tasks at least query units and counts explicitly, and the foils were chosen to limit the use of alternative strategies to the extent possible. Furthermore, although these tasks have been used less frequently than the approximate tasks, when they have been used, they are more likely to reveal poor performance and conceptual weaknesses, suggesting that they may prevent responding based on partial or approximate knowledge. Critically, the present study is the first to compare individual children's *developing* abilities in these syntactic and approximate tasks. The Base-10 counting task directly assesses the count+unit structure (e.g., Chan et al., 2014; Fuson, 1990; Kamii, 1986) by asking children to count sets grouped into base-10 units, to count, for example, 1 set of 100, 3 sets of ten, and 2 sets of 1. Chan et al. (2014) found that counting behavior and accuracy in determining the total number were both predictive of later mathematical achievement. Expanded notation directly assesses the underlying place values of written notation and asks children to match written numerals to their expanded notation forms (e.g., "324" = 300 + 20 + 4, Barrouillet et al., 2004; Mix et al., 2016). Finally, the "Which *N* has __?" task asks children directly about the count of base-10 units, as in "Which number has 3 tens?" Foils were chosen to limit heuristics-based reasoning, (e.g., providing options such as "34," "243," and "342" for the question, "Which number has 3 hundreds?"). Thus, unlike the approximate tasks, the three syntactic tasks all directly asked children about base-10 units and their counts.

Current study

Children were first assessed in the spring of kindergarten, an age at which children have exhibited early but partial understanding of base-10 symbols in previous work (e.g., Byrge et al., 2014; Mix et al., 2014; Yuan, Smith et al., 2019). The children were assessed a second time in the spring of first grade, at an age when they typically have been exposed to base-10 symbols in school instruction, and also at the age when previous research tells us both that (1) children vary in their syntactic knowledge and (2) this variability is highly predictive of later mathematics performance (e.g., Moeller et al., 2011). At both timepoints, children were assessed with the three approximate and three syntactic tasks. In first grade, children were also given a measure of general cognitive ability (Matrix Reasoning, WISC-V).

If these tasks, as we have grouped them, reflect different underlying knowledge structures, then we should see evidence for separate latent constructs in our factor analysis. Furthermore, if early approximate knowledge prepares children in some way for learning about count+unit syntax, then kindergarten children who perform well on approximate tasks should learn more about count+unit syntax in first grade and perform better on syntactic tasks than children who performed less well on approximate tasks. If approximate knowledge can hide deficits in emerging knowledge of count+unit syntax, then there should be children at the end of first grade with strong approximate knowledge—and who if we used these tasks as measures would be assessed as competent—yet nonetheless have a poor understanding of count+unit syntax.

METHOD

Participants

At the first test session, conducted between March and June of 2017, a total of 279 kindergartners (135 females; 144 males) with a mean age of 5.76 years (SD = 0.55) participated. At the second test session, conducted between March and June of 2018, when the children were students in first grade, 232 of them were assessed. The remaining 47 children had moved away from the participating schools and could not be located. Our target sample size was 120 children, based on the recommendation of including 20 children per measure for a confirmatory factor analysis (CFA; MacCallum et al., 1999; Raykov & Marcoulides, 2010), and six measures total. A sample size of 120 is also adequate to detect a medium effect in the multiple regression models (i.e., Cohen's f^2 = .11, Cohen, 1988), according to a sensitivity test conducted in G*Power with alpha of .05; power of .80; sample size of 120; and 6 predictors (Faul et al., 2009). Thus, with 232 participants, the study had more than adequate power. Note that we also carried out a latent variable path analysis, and though this analysis was not included in our a priori sample size planning, the analysis is comparable to multiple regression with manifest variables and tends to require fewer subjects than its corresponding regression model(s) to detect specific effects, so we are confident that our final sample size was adequate for this analysis as well.

Children were recruited from the Midwest and Mid-Atlantic regions of the United States. There were 40 children from Indiana, 186 children from Maryland, and 53 children from Michigan. We obtained written consent for children's participation from parents at all but one site where families were provided with an opt-out consent form at the school's request. Overall, 46% of the sample reported demographic data—the missing parent-report data were primarily from one site due to the use of an opt-out consent process. To account for this missing parent-report data, we used school-wide demographic information for three schools within this site. The fourth school within this site, however, did not have available school-wide demographic information and instead, we used this school's 2017 neighborhood census data. Once we had data for each site and school, we computed overall, weighted sample descriptive statistics, indicating that the sample was drawn from an ethnically diverse (41% Black, 38% White, 8% Asian; 12% Latino) population, and primarily middle socioeconomic status (average median family income range = $75,000–$99,999).

Procedure and materials

Children were tested twice—once in the spring of kindergarten and again in the spring of first grade. Testing sessions took place in a quiet area outside the classroom and lasted approximately 60 min per child. All measures were administered individually in one of two random orders, pseudo-randomized and counterbalanced across children—see Supporting Information for details on these task orders, items for each task, and mean kindergarten and first-grade performance for each item across all tasks. Reliabilities were calculated at both time points using Cronbach's α. Certain measures had low levels of internal consistency (i.e., reliabilities below .70), which might suggest multidimensionality within the task and weaken the strength and generalizability of the results in the current study. However, it has been argued that such measures may be retained if they provide important content coverage (Schmitt, 1996). As an added assurance, we computed an alternative reliability metric to assess replicability of the latent constructs—Coefficient *H*—that measures maximal reliability for an optimally weighted scale (Hancock & Mueller, 2001). These metrics indicate high construct replicability of the latent variables vis-à-vis their task indicators. The measures are described in detail below.

Approximate place value measures

In the *magnitude comparison* task (Mix et al., 2014), children were asked which of two written numerals represented the larger quantity (e.g., 461 vs. 614). The choice numerals were printed on the right and left sides of an 8 × 11-in. sheet of paper and trials were presented one by one by flipping pages in a binder. There were 25 trials comprised of 1- to 4-digit numerals (see Table S2). The comparisons were adapted from Mix et al. (2014). Correct responses received one point, for a total possible of 25 (chance = 12; α = .72 at Time 1 and .79 at Time 2).

In the *number line estimation* task (Siegler & Opfer, 2003), children were presented with a blank 0–100 number line and told to indicate where a number (e.g., 3) should be located using a vertical hash mark. There was one practice trial and 15 test trials. The test trials were coded for percentage of absolute error by measuring the distance from the hash mark to the correct location. Scores were then transformed by subtracting each score from 100 so that higher scores indicate better performance. The total score was the percent accuracy averaged across the 15 test trials (see Table S2; range = 0%–100%, even-odd reliability at Time 1: r = .76, and even-odd reliability at Time 2: r = .74).

In the *transcoding* task (e.g., Byrge et al., 2014), children's understanding of the mapping between spoken multi-digit names and written notation was assessed in both reading and writing tasks. For the reading assessment, children saw a stimulus number (e.g., "23") and said its name aloud while the experimenter recorded their response (e.g., "twenty-three"). For the writing assessment, children listened to the experimenter say a multidigit number name and were told to write down the numeral they heard. Both the reading and writing assessments were comprised of two 2-digit numbers; two 3-digit numbers; and two 4-digit numbers, for a total of 12 test trials across the two (see Table S2). Trials were coded as either correct or incorrect (maximum possible score = 12). Partially correct responses were not counted as correct (e.g., reading the numeral 239 as two hundred three-nine); however, full credit was given for written responses that involved numeral reversals (e.g., writing 3 backwards; α = .87 at Time 1 and .87 at Time 2).

Syntactic place value measures

In the *Base-10 counting* task (Chan et al., 2014), children were asked to count the number of squares in a simple line drawing of various quantities represented with base-10 blocks (see Figure 1). The critical question is whether children group the quantities by base-10 units as an aid to counting. That is, do they count the squares by base-10 units (e.g., counting 42: 10–20–30–40–41–42)? Or do they make errors such as treating all the units as ones (e.g., counting 42: 1–2–3–4–5–6) or counting only the ones blocks and ignoring the rest? To encourage counting by base-10 units, Chan et al. chose quantities that were too high to be counted easily by ones, and presented

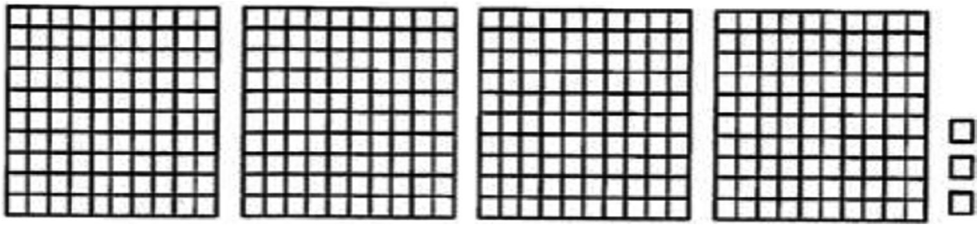

FIGURE 1 Sample item from Chan et al. (2014) base-10 counting task

line drawings that were so small it would be difficult to tag each of the individual small squares accurately (see Tables S2 and S3). We found in pilot-testing that children became easily fatigued and frustrated if they were allowed to attempt counting by ones, particularly at the kindergarten timepoint. Therefore, we took additional steps to discourage counting by ones, described below.

Prior to the test trials, children were first provided with a short 5-min introduction to three-dimensional base-10 blocks to ensure they understood what objects were depicted in the line drawings. The experimenter displayed a tens block and demonstrated how ten of the "small squares" could be combined to make a tens block, counting the individual squares by ones. Next, the experimenter introduced the hundreds block, first showing how 100 small squares came together to make the larger hundreds block. Initially, the experimenter counted the small squares by ones, but then stopped, noting that counting "all those small squares would take a really long time," showing instead how ten of the tens blocks could make the larger hundreds block, counting them by tens to illustrate. After this short introduction, children were presented with the first test trial—a line drawing of a physical representation of a quantity (e.g., 13 represented with base-10 blocks) and told it was a picture of the same blocks. Children were then asked to tell the experimenter how many small squares were in the picture.

The ten trials were presented in numerical order from smallest to largest (see Table S2). Children were permitted to count by ones on the first trial, but if they attempted to do so on the second trial (for which the target number was 42), they were allowed to finish, and then reminded that these blocks could also be counted by tens. Children were then allowed to count again, and the better of the two trials was scored. A similar prompt was given if children attempted to count the first trial with hundreds blocks by ones, but children were reminded they could count by hundreds. To avoid fatigue, children were stopped after the first hundreds trial if they had failed to count by base-10 units on previous tens items and continued to do so on the first hundreds item.

Chan et al. found that among first-grade students, accuracy was a reliable proxy for counting by base-10 units, so following their method, each of the 10 trials were coded as either correct or incorrect based on the child's final response of total number of blocks. Thus, each child's score was the total number of correct trials (maximum possible score was 10; α = .85 at Time 1 and .82 at Time 2).[1]

The *Which N has __?* task was a multiple-choice adaptation of the digit correspondence task (e.g., Hanich et al., 2001; Kamii, 1989). Children were presented with three written numerals arranged in a horizontal line (e.g., 2, 20, and 10). The experimenter then asked the child to select the number that answered a place value question such as, "Which number has two tens?" Prior to the test trials, there were two practice trials where the experimenter provided the child with corrective feedback. The six test trials included two trials each probing tens, hundreds, and thousands (see Table S2). The position of the correct response was counterbalanced across trials. Test trials were coded as either correct or incorrect (maximum possible score = 6, chance = 2; α = .53 at Time 1 and .59 at Time 2).

The *expanded notation* task is a commonly used multiple choice task asks children to match written numerals to their expanded notation forms and has been used in previous research (e.g., Mix et al., 2016). Children were shown a written numeral (e.g., 11) and asked to select the correct expanded version from among three options (e.g., 10+1, 10+10, or 1+1). The choices were arranged vertically on the right side of the page, and the target number was presented in a larger font on the left side. Before the test trials, the experimenter explained that the plus sign means combining two numbers and illustrated this with a small set, such as 1 + 1 = 2. Then they asked, "Which of these (pointing to the equations) adds up to be this number (pointing to the target?)." The goal of the task was to select the two addends that when added together made the target number. Children completed two practice trials for which corrective feedback was offered. The six test trials included two trials each probing 2-digit, 3-digit, and 4-digit numbers (see Table S2). Test trials were coded as either correct or incorrect, for a maximum possible score of 6 (chance = 2; α = .66 at Time 1 and .70 at Time 2).

[1]By this scoring method, it would be possible for children to obtain correct answers after counting by 1 s. An examination of our field notes revealed that this was the case on 7.92% of trials in kindergarten and 2.37% of trials in first grade. Most of these trials (61.59% in kindergarten and 81.48% in first grade) occurred on the first item which was also the lowest number queried (i.e., "13"). To ensure that these responses were not biasing our results, we repeated all analyses excluding trials where children were correct after counting by 1 s. The results in all cases were the same as when we used the original total correct variable, so we reported our original analyses and results in the main text.

General cognitive ability

To estimate general cognitive ability, we used children's scores on the Matrix Reasoning subtest from the Wechsler Intelligence Scale for Children, 5th edition (WISC–V; Wechsler, 2014). Matrix Reasoning was assessed at the second test session when children were in first grade. The subtest consists of two practice items and 32 test items in which children choose a figure that completes a repeating pattern or visual analogy. Based on the WISC–V testing procedures, children completed items that increased in difficulty until they gave three consecutive incorrect responses. Age-based standardized scores were used as a covariate in all reported analyses. Reliability from the norming sample was high ($\alpha > .80$).

Analysis approach

We first assessed our a priori characterization of the tasks as assessing separable classes of emerging knowledge: approximate and syntactic. To this end, we conducted separate independent CFAs at kindergarten and first grade, asking specifically whether the hypothesized 2-factor solution was better than a 1-factor solution. For this and all remaining analyses, raw scores (i.e., total correct) for each task were used, however, reported estimates are from the standardized solutions. Chi-square tests were used to compare the fit between the two solutions at each time point. A residualized dataset ($n = 230$) was created for these analyses by regressing scores from each task onto Matrix Reasoning performance, thereby controlling for performance in general cognitive ability. We also used a latent variable path analysis to assess whether performance on one of the factors predicted performance on the other.

To investigate performance differences between approximate and syntactic measures, we carried out analyses of covariance across timepoints (kindergarten and first grade), with Matrix Reasoning scores as the covariate. We also examined group differences by dividing children into groups based on their performance on approximate tasks. To prepare the data for these analyses, we transformed children's raw scores on each of the six experimental tasks to proportions correct as follows. For transcoding and base-10 counting, we used children's proportions correct. For number line estimation, proportion correct was computed by dividing the reverse score of percent absolute error by 100. For the three forced-choice tasks (e.g., which number has __?, expanded notation; and magnitude comparison), chance-level responding or guessing was accounted for by subtracting the total number of incorrect items times $1/k - 1$ (where k is the number of answer choices) from the total number of correct items. The corrected score was then divided by the total number of items to produce a corrected proportion correct. For example, because expanded notation had six items with three answer choices, we subtracted the (number incorrect) × (1/(3 – 1)) from the total number correct. We then divided the corrected score by six to obtain a corrected proportion correct.

All reported factor and path analyses used maximum likelihood estimation with robust standard errors (i.e., MLR), which helps to protect against specification errors due to non-normal distributions (e.g., Chou & Bentler, 1995). For all analyses, models were assessed using data-model fit information, including the root mean square error of approximate (RMSEA) and comparative fit index (CFI) for models without a mean structure (e.g., CFAs). All analyses were conducted using Mplus software (Muthén & Muthén, 1998–2017). In each model, full information maximum likelihood estimation was used to accommodate missing data with rescaling corrections to standard errors and fit indices to handle potential nonnormality in the data. Reported estimates are from the standardized solutions.

RESULTS

Descriptive statistics

Children's mean performance on the six place value measures is presented in Table 1 and we provide visualizations of the distributions of performance as violin plots in Figure 2. The violin plots illustrate the distribution of performance across all participants for each time point. Wider sections of the violin plot represent a higher probability that participants of the population will take on the given value whereas skinnier sections represent a lower probability. As shown in these plots, children generally improved their performance on all six tasks across the period of the study, with many children reaching ceiling performance in first grade.

Individual children's performance on the two hypothesized factors (i.e., approximate and syntactic) at each timepoint is presented in Figure 3. As suggested by children's raw scores on individual tasks (Table 1), performance on the approximate tasks was better than performance on syntactic tasks at both timepoints, but there was considerable growth in both factors from kindergarten to first grade. Still, children's scores on individual measures, as well as scores for the two hypothesized factors (i.e., approximate and syntactic), were significantly intercorrelated (see Table 2), suggesting overlap in the competencies tapped, at least for some children. Note that the interfactor correlations were lower in kindergarten than in first grade.

Are approximate and syntactic tasks really distinct?

The basic premise of the present study is that the understandings underlying performance on approximate

TABLE 1 Means (*SD*s) on place value measures in kindergarten and first grade

Assessment	Kindergarten (*n* = 279) Mean number correct (*SD*)	Kindergarten (*n* = 279) Mean proportion correct (*SD*)	First grade (*n* = 231) Mean number correct (*SD*)	First grade (*n* = 231) Mean proportion correct (*SD*)
Magnitude comparison	19.07 (4.03)	.53 (.32)	22.67 (2.92)	.81 (.23)
Number line estimation	82.09 (8.75)	.82 (.09)	88.65 (7.44)	.89 (.07)
Transcoding	6.43 (3.08)	.54 (.26)	9.42 (2.89)	.78 (.24)
Base-10 counting	3.71 (2.77)	.37 (.28)	6.40 (2.68)	.64 (.27)
Which number has __?	4.60 (1.17)	.65 (.29)	5.01 (1.27)	.75 (.32)
Expanded notation	3.01 (1.78)	.25 (.45)	4.81 (1.48)	.70 (.37)

Note: Mean proportion correct scores of expanded notation, Which number has __?, and magnitude comparison are adjusted for forced-choice responding. The mean number correct for number line estimation was derived from the percent absolute error (PAE) score by subtracting the PAE from 100. The mean proportion correct score for number line estimation was derived by dividing this reversed PAE score by 100.

place value measures may differ from those underlying performance on syntactic place value measures. To test this idea, we conducted two CFAs—one at kindergarten and one at first grade—with the six tasks divided into two groups: (1) approximate tasks including magnitude comparison, number line estimation, and transcoding, and (2) syntactic tasks including base-10 counting, which number has__?, and expanded notation. As before, children's place value scores were residualized to control for differences in general cognitive ability, based on their Matrix Reasoning scores. The fit indices of the kindergarten 1-factor solution were CFI = .970 and RMSEA = .07, and the fit indices of the Grade 1-factor solution were CFI = .997 and RMSEA = .02. Additionally, the two-factor model fit was good when assessed separately for both kindergarten (CFI = .987; RMSEA = .05) and first grade (CFI = .993; RMSEA = .04); see Table 3. Furthermore, the place value measures hypothesized to tap syntactic understanding loaded significantly onto one factor and the place value measures hypothesized to tap approximate understanding loaded significantly onto the other factor. To assess reliability of the latent constructs, we calculated Coefficient *H* that measures maximal reliability for an optimally weighted scale (Hancock & Mueller, 2001). In kindergarten, both the syntactic and approximate factors' construct replicability were acceptable: $H_{\text{syntactic}}$ = .72, 95% CI [.66, .94] and $H_{\text{approximate}}$ = .72, 95% CI [.67, .80]. In first grade, both the syntactic and approximate factors' construct replicability were again acceptable: $H_{\text{syntactic}}$ = .74, 95% CI [.68, .85] and $H_{\text{approximate}}$ = .80, 95% CI [.75, .94]. These values indicate high construct replicability of the latent variables vis-à-vis their task indicators.

Chi-square tests revealed that the two-factor solution had statistically significantly better fit than the one-factor solution only in kindergarten ($\chi^2_{\text{diff}} = 8.97$, $p = .001$). In first grade, there was no detectable difference ($\chi^2_{\text{diff}} = 0.00$, $p = .959$). These findings suggest that approximate measures and syntactic measures plausibly assess separate, albeit highly correlated, latent constructs among kindergarten students, but by first grade, all six tasks appear to measure the same underlying understanding. Thus, as predicted, approximate and syntactic understanding of place value appear to be distinct, but only in kindergarten.

Do early informal ideas support the acquisition of syntactic knowledge?

Knowing that the approximate and syntactic factors were separable in kindergarten, we next asked whether performance on one or the other factor was more predictive of first-grade skill. One possibility is that early performance on syntactic tasks is the only significant predictor of first-grade performance on either approximate or syntactic tasks because decomposition skill contributes to both. On this account, early performance on approximate tasks may be a distraction that is unrelated to later performance if it is based on shortcuts and heuristics, such as guessing that longer numerals represent larger quantities. Alternatively, early approximate performance may be a significant predictor if it is based on partial knowledge that children can leverage to understand count+unit syntax (i.e., using the insight that longer numerals represent larger quantities to bootstrap into a more specific understanding of hundreds vs. tens, for example). Because the chi-square difference test suggested a single place value factor in first grade, we first conducted a CFA that examined the fit of a two-factor solution in kindergarten and one-factor solution in first grade (i.e., 2 × 1 model). Results indicated acceptable fit (CFI = .923, RMSEA = .08). Between the two paths tested, only latent approximate scores in kindergarten predicted place value understanding in first grade, $\beta = 1.25$, $p = .005$; syntactic scores in kindergarten did not, $p = .279$ (see Figure 4a). Thus, there is strong evidence that children's early heuristics and approximate understandings contribute positively to the eventual acquisition of syntactic understanding.

One might argue that the previous finding was due to early approximate skill predicting later approximate

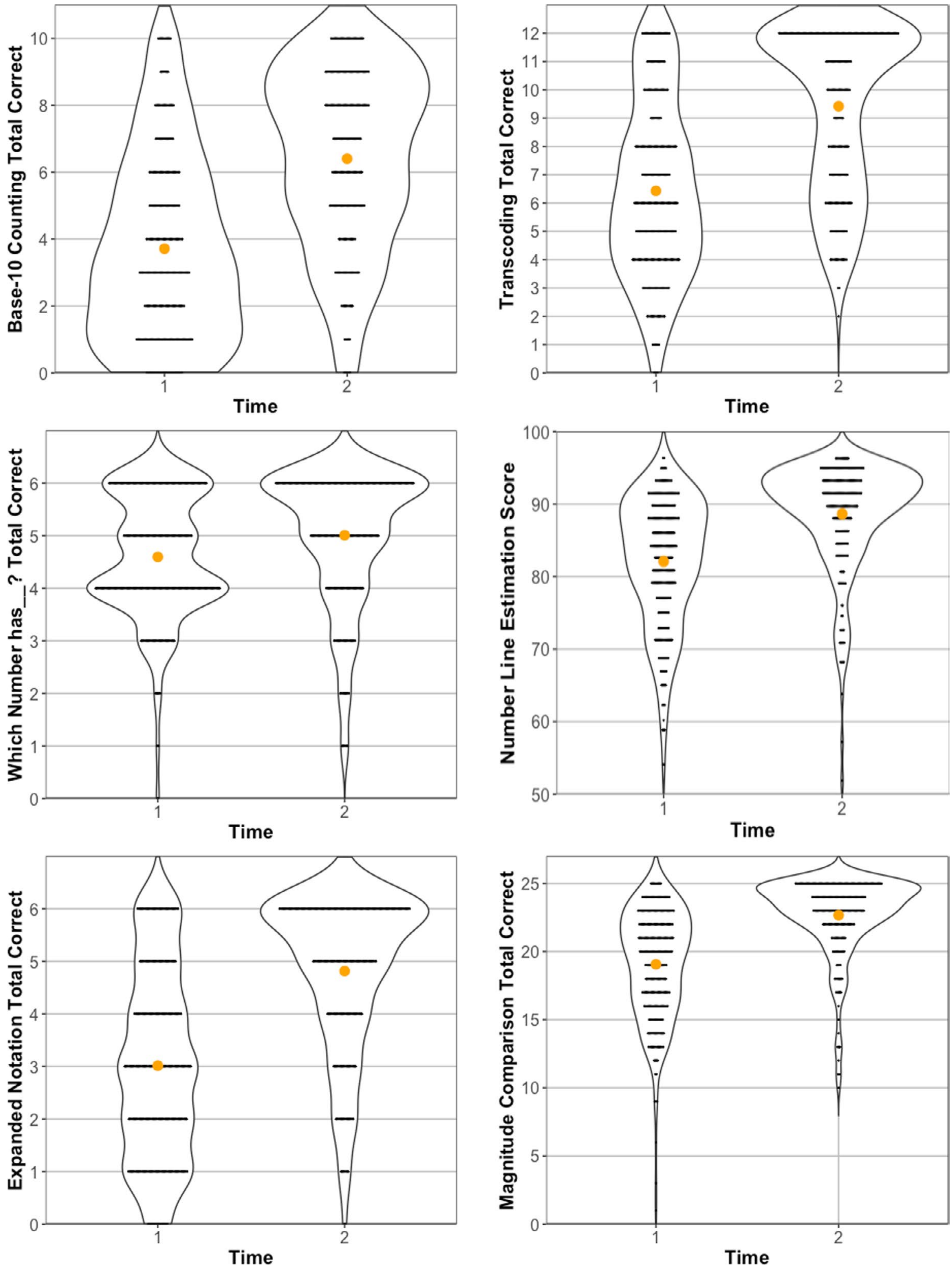

FIGURE 2 Violin plots for each task illustrate the distribution of performance across all participants for each time point. Wider sections of the violin plot represent a higher probability that participants of the population will take on the given value whereas skinnier sections represent a lower probability. The orange dot is the mean score for each time point

skill to such an extent that it appeared to predict both syntactic and approximate skill equally. To assess this possibility, we conducted a second path analysis that assumed both factors in both grades (i.e., a 2 × 2 model). The fit for this 2 × 2 model was poor (CFI = .861, RMSEA = .11; Hu & Bentler, 1999). Interestingly, though

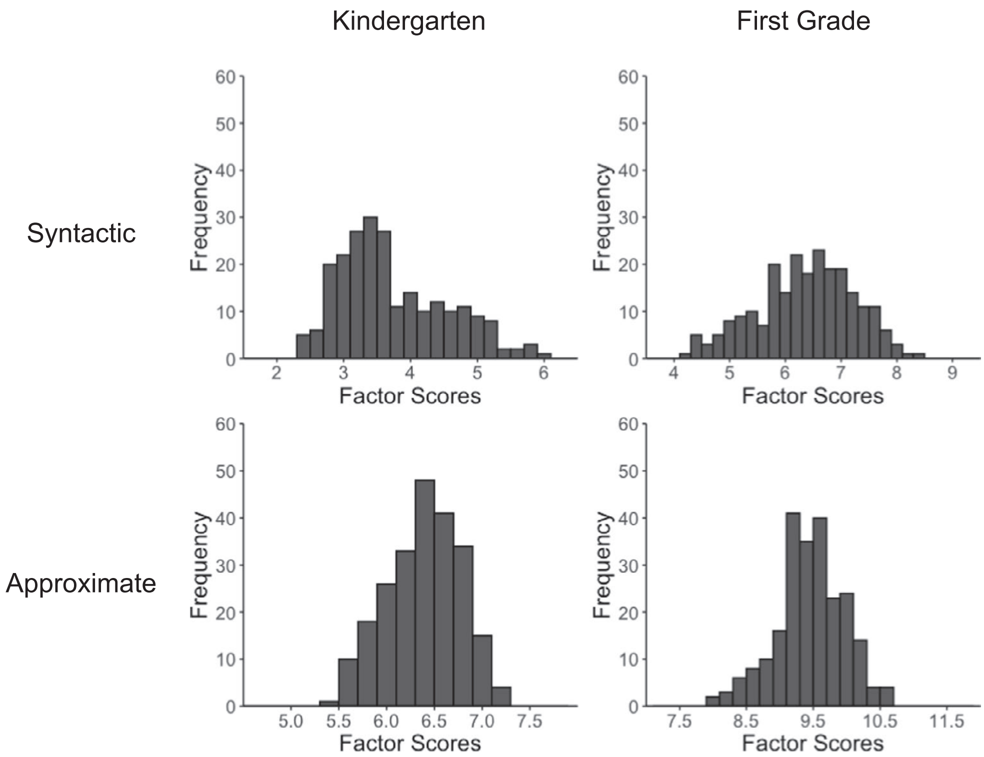

FIGURE 3 Histograms of the syntactic and approximate factor scores using the residualized dataset (controlling for Matrix Reasoning) at each time point

kindergarten performance on approximate tasks significantly predicted performance on both approximate, $\beta = 1.23$, $p = .001$, and syntactic tasks, $\beta = 0.91$, $p = .013$ in first grade, kindergarten performance on syntactic tasks predicted neither. These findings demonstrate that the predictive relations in the previous analysis were not limited to approximate tasks at first grade (see Figure 4b).

Does competence on approximate tasks mask deficiencies in count+unit understanding?

An inspection of children's mean performance on approximate and syntactic tasks (Figure 3) suggests their scores were lower on syntactic tasks. This posthoc observed pattern was confirmed in an exploratory repeated measures analysis of covariance (ANCOVA) with grade (kindergarten vs. first grade) and task type (approximate vs. syntactic) as within-subject variables, and Matrix Reasoning scores as the covariate. Specifically, there was a statistically significant main effect of task type ($F(1, 228) = 119.27$, $p < .001$, $\eta_p^2 = .34$), such that performance on approximate tasks ($M = 0.73$, $SE = .01$) was greater than on syntactic tasks ($M = 0.56$, $SE = .01$; see Figure 5). There also was a significant main effect of grade ($F(1, 228) = 34.66$, $p < .001$, $\eta_p^2 = .13$), such that scores on both task types improved from kindergarten to first grade; however, the interaction between grade and task type was not significant ($p = .370$). These findings relate to our hypothesis that children may appear to have more place value knowledge than they actually have, depending on how place value knowledge is measured. If children perform significantly better on tasks that admit the use of partial knowledge, approximate quantification, or heuristics than they do on tasks that require analysis of count+unit syntax, they may appear to know more about place value than they actually do.

To further investigate this idea, we examined children's performance on syntactic tasks after dividing them into groups based on their performance on approximate tasks. The low approximate group performed at or below the 33rd percentile ($n = 77$), the high approximate group performed at or above the 66th percentile ($n = 77$), and the moderate approximate group had percentile scores in between ($n = 77$). The mean proportion correct on syntactic tasks for children in each of these three groups is presented in Figure 6.

As one might expect, children in the high approximate group also performed relatively well on syntactic tasks and children who performed poorly on approximate tasks also performed relatively poorly on syntactic tasks. However, children who performed moderately on approximate tasks in kindergarten also performed quite poorly on syntactic tasks. To assess this posthoc observed pattern, an exploratory repeated measures ANCOVA with approximate performance group as the between-subject factor, grade as the within-subject factor, and the syntactic proportion correct composite as the dependent variable revealed a statistically significant interaction between approximate performance group and grade ($F(2, 226) = 8.72$, $p < .001$, $\eta_p^2 = .07$). A

TABLE 2 Correlations of the place value measures and the hypothesized latent variables at kindergarten and first grade

	Kindergarten								First grade							
		Approximate				Syntactic				Approximate				Syntactic		
Variable	Approximate	MC	NLE	TC	Syntactic	B-10	WN?	EN	Approximate	MC	NLE	TC	Syntactic	B-10	WN?	EN
Kindergarten																
Approximate		.81	.79	.80	.60	.63	.35	.42	.66	.54	.52	.60	.58	.52	.45	.43
MC	.81		.45	.49	.42	.44	.23	.30	.48	.43	.37	.41	.41	.38	.34	.26
NLE	.81	.46		.44	.47	.47	.30	.32	.56	.44	.50	.46	.52	.45	.41	.39
TC	.83	.53	.51		.56	.61	.31	.39	.54	.42	.37	.56	.47	.42	.33	.39
Syntactic	.66	.46	.50	.65		.79	.71	.81	.49	.39	.36	.47	.54	.47	.40	.43
B-10	.68	.48	.51	.69	.83		.31	.54	.44	.33	.33	.46	.49	.50	.36	.33
WN?	.42	.29	.34	.41	.75	.41		.32	.33	.27	.26	.29	.34	.23	.27	.32
EN	.48	.33	.36	.48	.82	.58	.39		.36	.31	.24	.35	.41	.37	.28	.34
First grade																
Approximate	.71	.50	.60	.62	.58	.55	.41	.43		.86	.81	.82	.74	.66	.55	.58
MC	.59	.45	.48	.49	.47	.41	.34	.37	.87		.56	.59	.62	.52	.49	.48
NLE	.58	.40	.54	.47	.46	.44	.35	.32	.84	.61		.47	.52	.51	.34	.41
TC	.66	.44	.52	.63	.56	.55	.37	.41	.86	.64	.56		.71	.62	.53	.57
Syntactic	.64	.44	.55	.57	.63	.59	.43	.47	.79	.67	.62	.76		.83	.78	.82
B-10	.59	.42	.49	.52	.57	.59	.34	.44	.73	.59	.60	.69	.86		.47	.54
WN?	.53	.38	.45	.45	.50	.48	.36	.36	.63	.55	.46	.61	.82	.56		.42
EN	.50	.30	.43	.47	.51	.43	.39	.40	.65	.54	.50	.63	.84	.60	.51	

Note: Bivariate correlations are presented below the diagonal and partial correlations (controlling for Matrix Reasoning) are presented above the diagonal. Syntactic = composite of B-10, WN?, and EN tasks. Approximate = composite of MC, NLE, and TC tasks.

Abbreviations: B-10, base-10 counting task; EN, expanded notation; MC, magnitude comparison; NLE, number line estimation; TC, transcoding; WN?, Which number has __?

All correlations were significant at $p < .05$.

TABLE 3 Latent factor loadings by grade

	Kindergarten		First grade	
Assessment	Approximate	Syntactic	Approximate	Syntactic
Magnitude comparison	.612		.742	
Number line estimation	.623		.641	
Transcoding	.775		.810	
B-10 counting		.864		.751
Which number has __?		.394		.628
Expanded notation		.620		.685

Note: All factor loadings are significant, $p < .05$. The estimates presented here are from the standardized solution, which is why loading constraints are not shown.

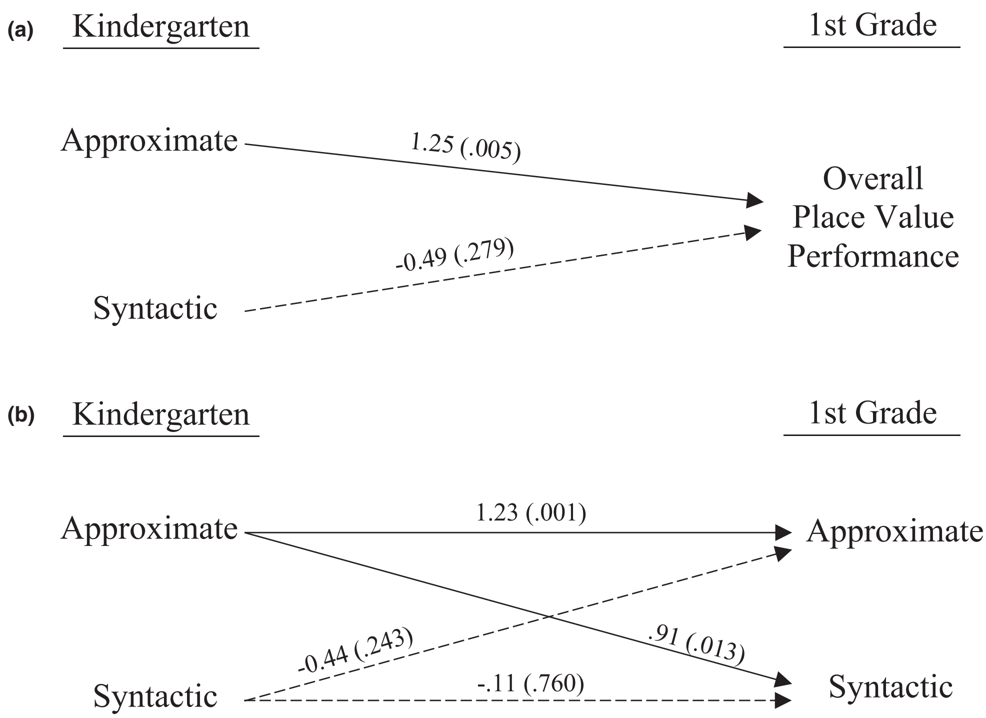

FIGURE 4 Latent variable path analysis of kindergarten approximate and syntactic latent variables of place-value understanding predicting first-grade variables using the residualized dataset: (a) 2 × 1 factor structure and, (b) 2 × 2 factor structure. Standardized coefficients (β) are presented with p-values in parentheses. Although covariances between approximate and syntactic variables at both kindergarten (in a and b) and first grade (b only) were included in the model, they are not shown in this figure

simple effects test suggests this is likely because the difference in syntactic scores in kindergarten was nonsignificant between the low and moderate approximate groups ($M_{diff} = 0.06$, $SE_{diff} = .03$, $p = .052$), whereas in first grade the difference between the low and moderate groups was significant ($M_{diff} = 0.15$, $SE_{diff} = .03$, $p < .001$). This finding lends further support to our hypothesis that children who show some understanding of place value on approximate tasks may nonetheless perform quite poorly on syntactic tasks. Thus, although early approximate understanding may be an important foundation for later syntactic understanding, the two understandings are not interchangeable, at least early in development.

DISCUSSION

This study addressed fundamental questions about how children build their understanding of complex representational systems. The basic challenge facing learners is how to discover the underlying structure of these systems based on the fragments, false starts, and rough approximations that likely characterize initial attempts at sense-making. Do these early but imperfect understandings pave the way for later mastery? Or are they detours leading nowhere, perhaps even becoming obstacles to discovering relational structure? For the current study, we addressed these questions in the context of place value—a particularly challenging complex system for

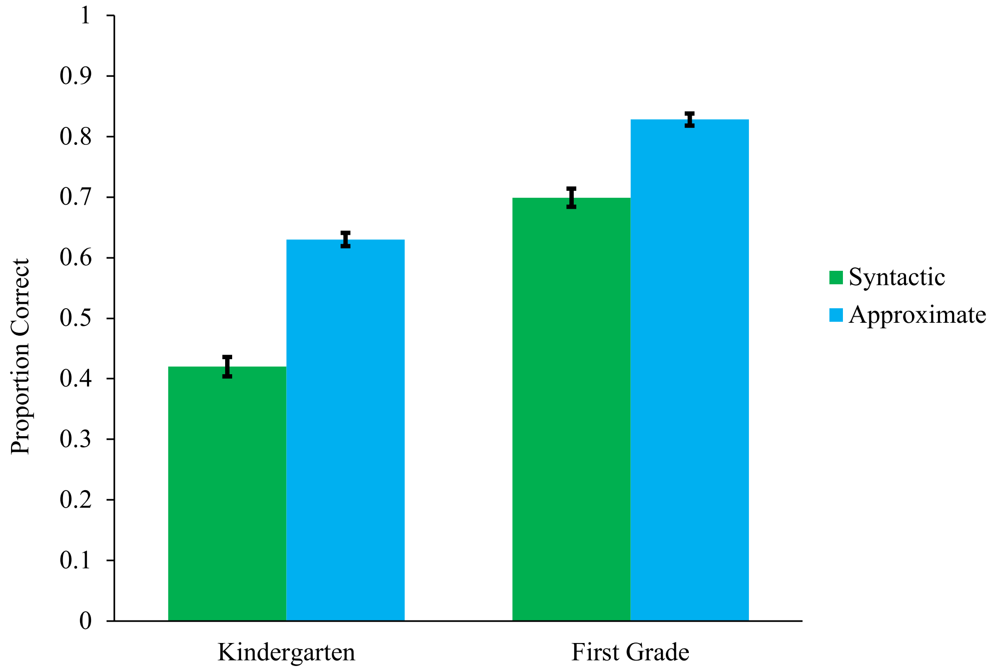

FIGURE 5 Bar graphs display estimated marginal means of proportion correct for both approximate and syntactic place value understanding by grade, controlling for general cognitive ability. Error bars indicate standard error

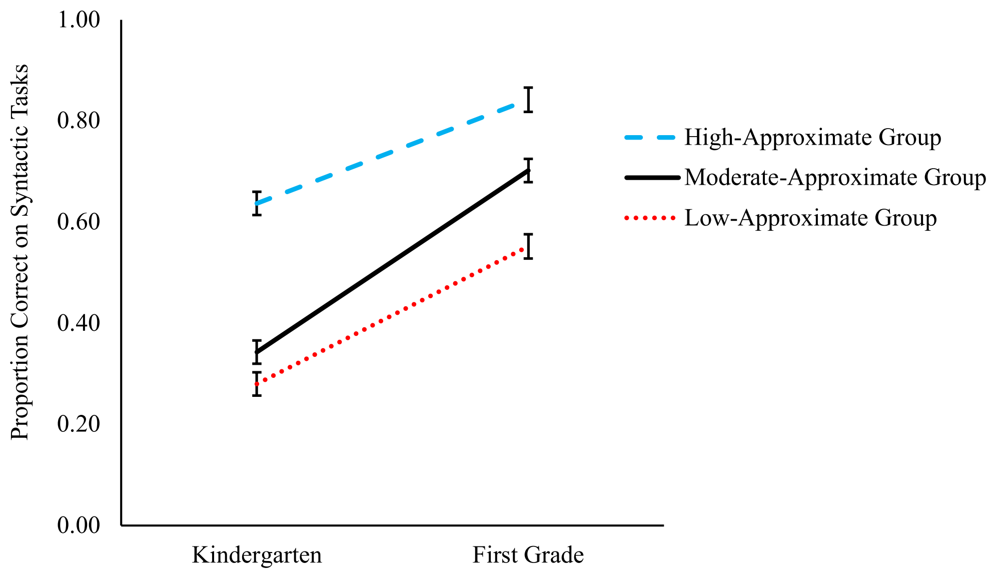

FIGURE 6 The line graph displays estimated marginal means of the syntactic tasks' composite proportion correct across time by approximate performance group, controlling for general cognitive ability. Error bars indicate standard error.

which early deficits in understanding are subsequently linked to poor mathematics outcomes (Chan et al., 2014; Moeller et al., 2011; Ross, 1986). Specifically, we assessed children's place value understanding using the same six measures (three approximate and three syntactic) at two timepoints (kindergarten and first grade), to see how performance on these tasks related across development.

Before we could address our developmental hypotheses, we first had to confront a difficult measurement problem. Based on theory, we hypothesized that performance on place value tasks that require explicit understanding of count+unit syntax tap different latent constructs than tasks that allow for responses based on approximate understanding, and thus may contribute differentially to development. A CFA yielded support for this idea, demonstrating that approximate and syntactic place value understanding are distinct (albeit correlated) constructs in kindergarten. By first grade, these differences had disappeared and performance on the six tasks converged. Perhaps as children acquire count+unit syntax, they apply this syntactic understanding to all place value tasks, leading to a convergence of the two systems (approximate and syntactic) by first grade. That said, syntactic tasks were more difficult than approximate tasks, and this difference in difficulty was still apparent in first grade, among children at all levels of achievement. Thus, while there may be some conceptual convergence in this age range, it appears that children still perform better on the less syntactically demanding, approximate measures.

Interestingly, whereas one might expect early understanding of count+unit syntax to predict later achievement, that was not the case. Path analyses showed that *only* performance on the approximate tasks in kindergarten significantly predicted first-grade place value performance, including performance on syntactic measures. Based on these results, it appears approximate or partial understanding of place value symbols is an important precursor to the decomposition skill that has predicted later mathematics in previous studies. We speculated that children can perform approximate tasks above chance using simple heuristics, such as knowing longer numerals are larger than shorter numerals in the magnitude comparison task, or rough mappings, such as mapping the temporal order of words and suffixes in spoken number names to the left-right order of written digits in the transcoding task. Perhaps, these early piecemeal or rough understandings of place value orient attention to places in written numerals and units in number names, eventually leading to a fine-tuning as more specific place meanings are discovered. Similar patterns have been demonstrated in the acquisition of small number meanings (e.g., Mix, 2009; Sarnecka & Carey, 2008) and acquisition of non-numeric symbols, such as color words (e.g., Sandhofer & Smith, 1999).

Despite the clear evidence of a key developmental role for approximate understanding, our results indicate that we cannot use approximate tasks to measure count+unit understanding, especially early in development, because these tasks measure different underlying competencies. We know from studies focused on measures that require knowledge of base-10 syntax that children who fail to understand decomposition in first and second grade have worse mathematics outcomes, and may have trouble understanding multidigit calculation and subsequent arithmetic operations (Chan et al., 2014; Moeller et al., 2011; Ross, 1986). Our results indicate that such deficits in children's place value understanding may be missed in kindergarten and first grade if only approximate tasks are used as measures.

It is interesting that, although children at all achievement levels showed significant growth in syntactic understanding across time, they neither reached ceiling nor fully converged by the end of first grade. Instead, all three performance groups showed improvement in both approximate understanding and syntactic understanding, and remained distinct even as they were acquiring base-10 understanding. A possible explanation for this developmental pattern is that children learn place value concepts via a stepwise bootstrapping process, perhaps one unit at a time. Previous research has suggested that 2-digit number meanings are acquired earlier and processed differently than 3-digit numbers as they are learned (Fuson, 1990; Kouba et al., 1988; Mann et al., 2012). If these performance differences extended to the present tasks, the worse performance of kindergarten children in the lower of the approximate performance groups may be due to approximate understanding for all three places we tested (tens, hundreds, thousands), leading children to obtain significantly higher scores on approximate tasks versus the syntactic tasks. As children in this group progress to first grade, they may acquire syntactic meanings for two-digit numbers first, leading to better (but not ceiling) performance on both the approximate and syntactic measures, and perhaps contributing to a unitary latent construct in first grade even though performance on the larger numbers remains distinct.

Now consider the same mechanism playing out in the higher performing approximate group. These children may have already acquired syntactic meanings for two-digit numbers by kindergarten, so their starting scores would be higher on both approximate and syntactic tasks compared to the lower performing groups. However, children in the higher performing group showed incremental growth similar to children in the lower performing group. On our stepwise interpretation, this pattern could be due to acquiring three-digit number meanings in first grade, but still not comprehending four-digit numbers or greater. Unfortunately, because many of the items on our tasks mixed the number of digits (e.g., comparing a two-digit number to a three-digit number, or choosing the expanded notation equivalent of a three-digit number from among choices that had various combinations of tens, hundreds, and thousands), it is not possible to test this interpretation directly in the present study, but this could be an interesting direction for future research.

In summary, our findings indicate that approximate and syntactic understandings of place value are distinct early in development, and eventually converge at the end of first grade, with a clear predictive relation from performance on approximate place value tasks to performance on place value overall in first grade. Children at all ability levels performed better on approximate than syntactic tasks, suggesting that there is incremental growth in both understandings across this age period. Though the present study was not designed to test this notion, our findings are consistent with a stepwise developmental bootstrapping process that could plausibly be based on learning one place (tens, hundreds, thousands, etc.) at a time.

These findings raise questions about the importance of approximate understanding of place value in developing these concepts and whether this understanding can be nurtured and perhaps leveraged by teachers. Indeed, little is known about children's informal exposure to multidigit numbers and large sets at home or in preschool Understanding the natural sources of variation in this exposure is an important next step for future research because this variation could explain why some children develop approximate skills earlier than others. Our findings also shed light on why children in the same classrooms vary in their eventual mastery of place value concepts. Clearly, children who exhibit strong performance on approximate tasks early in development, go on to have a better grasp of count+unit syntax in first grade. Perhaps by identifying children with weak approximate skills in kindergarten and first grade, teachers could intervene sooner to help them catch up.

However, some children (those in the low and moderate approximate groups) appeared slower to learn count+unit syntax, a deficit that may be masked if only approximate skills are assessed. Thus, children may also benefit from direct instruction in multidigit number meanings earlier than it is typically provided. That is, rather than waiting until first grade, teachers might consider introducing these concepts earlier, in kindergarten or even preschool. However, an important take-home message from our findings is that mastery of count+unit syntax in these early years need not be the goal. Rather, young children likely benefit from encouragement to first develop a rough or partial understanding, perhaps by simply increasing exposure to multidigit numbers or through activities that permit approximation, such as magnitude comparison or transcoding, building from rough to increasingly specific or nuanced comparisons.

ACKNOWLEDGMENTS

We thank the teachers and families of Wyoming, Michigan, Howard County, Maryland, Prince Georges County, Maryland, and Bloomington, Indiana, for their generous participation in this study. This research was supported by DRK-12 1664781 from the National Science Foundation to K. S. Mix and DRK-12 1621093 from the National Science Foundation to L. B. Smith. The data are published here: https://drum.lib.umd.edu/handle/1903/21495.

ORCID

Kelly S. Mix https://orcid.org/0000-0002-4402-2557
Corinne A. Bower https://orcid.org/0000-0001-7375-2842

REFERENCES

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004). ADAPT: A developmental, asemantic, and procedural model for transcoding from verbal to Arabic numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295x.111.2.368

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal*, *108*(2), 115–130. https://doi.org/10.1086/525550

Bugden, S., & Ansari, D. (2011). Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition*, *118*(1), 32–44. https://doi.org/10.1016/j.cognition.2010.09.005

Byrge, L., Smith, L. B., & Mix, K. S. (2014). Beginnings of place value: How preschoolers write three-digit numbers. *Child Development*, *85*(2), 437–443. https://doi.org/10.1111/cdev.12162

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Epson, S. (1997). A longitudinal study of invention and understanding in children's multi-digit addition and subtraction. *Journal for Research in Mathematics Education*, *29*, 3–20. https://doi.org/10.2307/749715

Chan, W. W. L., Au, T. K., & Tang, J. (2014). Strategic counting: A novel assessment of place-value understanding. *Learning and Instruction*, *29*, 78–94. https://doi.org/10.1016/j.learninstruc.2013.09.001

Cheung, P., & Ansari, D. (2021). Cracking the code of place value: The relationship between place and value takes years to master. *Developmental Psychology*, *57*(2), 227–240. https://doi.org/10.1037/dev0001145

Chou, C. P., & Bentler, P. M. (1995). Estimation and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

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*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fuson, K. C. (1988). *Children's counting and number concept*. Springer.

Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction*, *7*(4), 343–403. https://doi.org/10.1207/s1532690xci0704_4

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 293–304. https://doi.org/10.1177/00222194050380040301

Hancock, G.R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 195–216). Scientific Software International.

Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, *93*(3), 615–626. https://doi.org/10.1037/0022-0663.93.3.615

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. https://doi.org/10.1037/a0014939

Kamii, C. (1986). Place value: An explanation of its difficulty and educational implications for the primary grades. *Journal of Research in Childhood Education*, *1*(2), 75–86. https://doi.org/10.1080/02568548609594909

Kamii, C. (1989). *Young children continue to reinvent arithmetic, 2nd grade*. Teachers College Press.

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*(8), 14–19. https://doi.org/10.5951/AT.35.8.0014

Krajewski, K., & Schneider, W. (2009a). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*, *19*(6), 513–526. https://doi.org/10.1016/j.learninstruc.2008.10.002

Krajewski, K., & Schneider, W. (2009b). Exploring the impact of phonological awareness, visual–spatial working memory, and preschool quantity–number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*(4), 516–531. https://doi.org/10.1016/j.jecp.2009.03.009

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Mann, A., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H. C. (2012). On the development of Arabic three-digit number processing in primary school children. *Journal of Experimental Child Psychology*, *113*(4), 594–601. https://doi.org/10.1016/j.jecp.2012.08.002

Mazzocco, M. M., Myers, G. F., Thompson, L. A., & Desai, S. S. (2003). Possible explanations for children's literal interpretations of homonyms. *Journal of Child Language*, *30*(4), 879–904. https://doi.org/10.1017/S0305000903005786

Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research and Practice*, *20*(3), 142–155. https://doi.org/10.1111/j.1540-5826.2005.00129.x

Mix, K. S. (2009). How Spencer made number: First uses of the number words. *Journal of Experimental Child Psychology*, *102*(4), 427–444. https://doi.org/10.1016/j.jecp.2008.11.003

Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology: General*, *145*(9), 1206–1227. https://doi.org/10.1037/xge0000182

Mix, K. S., Prather, R. W., Smith, L. B., & Stockton, J. D. (2014). Young children's interpretation of multidigit number names: From emerging competence to mastery. *Child Development*, *85*(3), 1306–1319. https://doi.org/10.1111/cdev.12197

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities*, *32*(5), 1837–1851. https://doi.org/10.1016/j.ridd.2011.03.012

Morgan, P. L., Farkas, G., & Qiong, W. U. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, *42*(4), 306–321. https://doi.org/10.1177/0022219408331037

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(6), 491–505. https://doi.org/10.1002/wcs.1369

Odic, D., Lisboa, J. V., Eisinger, R., Olivera, M. G., Maiche, A., & Halberda, J. (2016). Approximate number and approximate time discrimination each correlate with school math abilities in young children. *Acta Psychologica*, *163*, 17–26. https://doi.org/10.1016/j.actpsy.2015.10.010

Piaget, J. (1952). *The origins of intelligence in children*. W W Norton & Co.

Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*, 237–254. https://doi.org/10.1080/01690969008402106

Raghubar, K., Cirino, P., Barnes, M., Ewing-Cobbs, L., Fletcher, J., & Fuchs, L. (2009). Errors in multi-digit arithmetic and behavioral inattention in children with math difficulties. *Journal of Learning Disabilities*, *42*(4), 356–371. https://doi.org/10.1177/0022219409335211

Raykov, T., & Marcoulides, G. A. (2010). Group comparisons in the presence of missing data using latent variable modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(1), 134–149. https://doi.org/10.1080/10705510903439136

Ross, S. (1986). *The development of children's place-value numeration concepts in grades two through five*. Presented at the annual meeting of the American Educational Research Association.

Ross, S., & Sunflower, E. (1995). *Place-Value: Problem-solving and written assessment using digit-correspondence tasks*. Presented at the annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, OH.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*(3), 668–679. https://doi.org/10.1037/0012-1649.35.3.668

Sarnecka, B., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674. https://doi.org/10.1016/j.cognition.2008.05.007

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350–353. https://doi.org/10.1037/1040-3590.8.4.350

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*, 281–300. https://doi.org/10.1111/j.2044-835X.1994.tb00635.x

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–243. https://doi.org/10.1111/1467-9280.02438

Wechsler, D. (2014). *WISC–V: Wechsler Intelligence Scale for Children integrated technical and interpretive manual* (5th ed.). Pearson.

Yuan, L., Prather, R. W., Mix, K. S., & Smith, L. B. (2019). Preschoolers and multi-digit numbers: A path to mathematics through the symbols themselves. *Cognition*, *189*, 89–104. https://doi.org/10.1016/j.cognition.2019.03.013

Yuan, L., Prather, R. W., Mix, K. S., & Smith, L. B. (2020). Number representations drive number-line estimates. *Child Development*, *91*(4), e952–e967. https://doi.org/10.1111/cdev.13333

Yuan, L., Smith, L. B., & Mix, K. S. (2019, October). Learning numbers as a system of symbols and their relations. In D. Kim's (Chair), *Symposium: The symbol grounding problem in numerical cognition: Insights from developmental psychology*, presented at the biennial meeting of the Cognitive Development Society, Louisville, KY.

Yuan, L., Xiang, V., Crandall, D., & Smith, L. (2020). Learning the generative principles of a symbol system from limited examples. *Cognition*, *200*, 1–15. https://doi.org/10.1016/j.cognition.2020.104243

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Mix, K. S., Bower, C. A., Hancock, G. R., Yuan, L., & Smith, L. B. (2022). The development of place value concepts: Approximation before principles. *Child Development*, 00, 1–16. https://doi.org/10.1111/cdev.13724