Published in final edited form as:

J Educ Psychol. 2012 August ; 104(3): 647–660. doi:10.1037/a0029018.

Building Kindergartners' Number Sense: A Randomized Controlled Study

Nancy C. Jordan, Joseph Glutting, Nancy Dyson, Brenna Hassinger-Das, and Casey Irwin
University of Delaware

Abstract

Math achievement in elementary school is mediated by performance and growth in number sense during kindergarten. The aim of the present study was to test the effectiveness of a targeted small group number sense intervention for high-risk kindergartners from low-income communities. Children were randomly assigned to one of three groups ($n = 44$ in each group): a number sense intervention group, a language intervention group, or a business as usual control group. Accounting for initial skill level in mathematical knowledge, children who received the number sense intervention performed better than controls at immediate post test, with meaningful effects on measures of number competencies and general math achievement. Many of the effects held eight weeks after the intervention was completed, suggesting that children internalized what they had learned. There were no differences between the language and control groups on any math-related measures.

Keywords

at-risk; mathematics; number sense; low income; intervention

Kindergarten achievement has far-reaching consequences. Kindergarten test scores are associated with college attendance, earning potential, and financial management, even when background characteristics are held constant (Chetty et al., 2010). Kindergarten mathematics, in particular, matters for long-term outcomes. Independent of cognitive ability and social class, kindergarten math concepts (e.g., knowledge of numbers and ordinality) are powerful predictors of adolescent learning outcomes across content areas (Duncan et al., 2007). Although most kindergarten math curricula cover multiple topics, number sense is of primary importance during this age period (National Research Council, 2009). Virtually all of the kindergarten mathematics topics in the Common Core State Standards (CCSS) (Common Core State Standards Initiative, 2010) are associated with number competencies related to knowledge of whole numbers, number relations, and number operations. Even measurement and geometry standards require children to use number words and concepts.

The goal of the present study was to refine and test a research-based number sense intervention on a population of kindergartners at risk for mathematics learning difficulties,

namely minority children attending schools in low-income communities. Low-income children are two times more likely to repeat a grade and 1.5 times more likely to be diagnosed with a learning disability than are middle-income children (Duncan & Brooks-Gunn, 2001). Children from low-income families show delays in mathematics relative to the norm as early as preschool (Clements & Sarama, 2008), a gap that does not attenuate during the school years (National Mathematics Advisory Panel, 2008). Minority African-American and Latino children are disproportionately represented among the lower social classes (Royer & Walles, 2007), resulting in large racial disparities in mathematics achievement.

The Need for Number Sense

Broadly speaking, number sense refers to understanding of number and operations, such as knowing that each number in the counting sequence is always one more than the one that comes right before it or one less than the number that comes right after (Jordan, Kaplan, Ramineni, & Locuniak, 2009; Schaeffer, Eggleston, & Scott, 1974; Siegler & Jenkins, 1989). Deficient number sense is a core marker for severe and persistent learning disabilities in mathematics (Mazzocco, Feigenson, & Halberda, 2011).

Foundational number sense can be viewed as nonverbal vs. symbolic (Cirino, 2011). Although nonverbal number sense appears to be foundational to learning the symbolic number system (Libertus, Feigenson, & Halberda, 2011) (see Rips, Bloomfield, and Asmuth (2008) for a counter argument), it is not sufficient for learning complex mathematics. Symbolic number sense reflects understanding of number words and written symbols and is strongly influenced by input and instruction (National Research Council, 2009). In preschool and kindergarten, children learn to apprehend the numerical values associated with small quantities automatically (i.e., verbal subitization) (Le Corre & Carey, 2007), but they use counting to determine the exact value of larger quantities (Baroody, 1987; Gelman & Gallistel, 1978). They also can compare the relative magnitudes of numerals (Case & Griffin, 1990) and perform simple arithmetic calculations (Ginsburg & Russell, 1981; Jordan, Huttenlocher, & Levine, 1992). Young children come to understand that each counting word describes a quantity that is one more than the previous one (n , $n + 1$, $(n + 1) + 1$, etc.) (Le Corre & Carey, 2007). More advanced number sense, acquired mainly through formal instruction, helps students make generalizations across numbers (National Mathematics Advisory Panel, 2008). Children grasp how numbers can be composed and decomposed, and develop principled understandings of place value (e.g., numbers 11 to 19 are made up of 10 and some ones) and the meaning of operations (e.g., $A + B = B + A$ for all numbers; if $A + B = C$ then $C - A = B$ or $C - B = A$) (Fuson, Grandau, & Sugiyama, 2001).

Early number competencies follow a developmental progression (Clements & Sarama, 2007). Most children move from core knowledge of numbers (e.g., small number recognition and counting), to number relations (e.g., number after knowledge, determining the larger of two numbers) to number operations (e.g., mentally adding and subtracting with small numbers) (National Research Council, 2009); the progression repeats as children learn larger numbers. Not surprisingly, foundational number sense in kindergarten is a strong predictor of math achievement in elementary school. The correlation between level of

number sense at kindergarten entry and general math achievement at the end of third grade is .63 (Jordan et al., 2009). Particularly predictive are the abilities to compare the magnitudes of numerals, to add and subtract small quantities, and to solve problems in different contexts. Kindergartners with strong number knowledge are more likely to use adaptive scaffolding strategies (e.g., counting on from an addend to solve an addition problem) and become fluent with addition and subtraction facts earlier than those with weaker number knowledge (Jordan, Kaplan, Ramineni, & Locuniak, 2008; Locuniak & Jordan, 2008). Similar findings have been observed in shorter-term longitudinal studies (i.e., kindergarten to first grade), which also show the importance of numeral identification and counting skill (Cirino, 2011; Clarke & Shinn, 2004; Lembke & Foegen, 2009) to mathematics growth.

Individual Variation in Number Sense

Despite relatively strong universal starting points associated with nonverbal number sense, individual differences in symbolic number sense emerge in preschool (Jordan et al., 1992; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; National Research Council, 2009). These differences are associated with previous experience and learning opportunities as well as intrinsic abilities. Low-income children enter kindergarten far behind middle-income children in number sense (Starkey, Klein, & Wakeley, 2004), and they are four times more likely than their middle-income counterparts to show flat growth on numeracy tasks during kindergarten and first grade (Jordan, Kaplan, Locuniak, & Ramineni, 2007; Jordan, Kaplan, Oláh, & Locuniak, 2006). Social class differences in preschool and kindergarten are larger on verbal number tasks, which do not use objects, than on nonverbal number tasks with object representations (e.g., Jordan et al., 1992). For example, low-income children perform worse than middle-income children on simple word problems and arithmetic combinations but not on nonverbal calculations that do not rely on number words or symbols. Although little research has been conducted on math development in young English language learners (ELL) (National Research Council, 2009), achievement differences favoring native language learners are evident by third grade (Abedi, 2004); effect sizes are largest on math subtests with the highest language demand. Bilingual children with a low command of the instructional language experience difficulties with mathematical word problems (Kempert, Saalbach, & Hardy, 2011). Dyscalculia, a math disorder with a known neurological basis, is characterized more by domain specific deficits in numerical knowledge than by general cognitive impairments (Butterworth & Reigosa, 2007). Students with dyscalculia seem to have particular difficulties with the symbolic system of number (Rousselle & Noël, 2007), although more fundamental problems with the universal magnitude system have been observed as well (Mazzocco et al., 2011).

Developing Number Sense through Targeted Instruction

There is evidence to suggest that early number sense can be developed through purposeful instruction (Griffin, 2004), although relatively few studies have conducted causal evaluations. At the preschool level, Clements and Sarama (2008) developed and tested the effectiveness of the Building Blocks mathematics curriculum. While the curriculum is comprehensive and covers a broad range of mathematical topics, number activities, such as

counting, number recognition, and number comparisons, are specifically taught in the 26-week instructional program. Using random assignment, pre- to post test scores showed that children who received Building Blocks instruction grew more than controls on a measure of early mathematical knowledge with meaningful effect sizes. Positive effects in numeracy also are seen in other comprehensive preschool math curricula (Dobbs, Doctoroff, Fisher, & Arnold, 2006; Klein & Starkey, 2008).

Using a more targeted approach, Ramani and Siegler (2008) randomly assigned low-income preschoolers in Head Start programs to play either a number-board game (1 to 10) or a contrast color board game that did not involve numbers. The number board game required children to spin a spinner and then move one or two numbers on the board until they reached 10. Children played the games for 20 minutes during four sessions over a two-week period. Relative to the children who played the color board game, children in the number board game condition made reliable gains from pre- to post test in their ability to compare which of two numerals is bigger, their ability to identify numerals by name, and their ability to count from 1 to 10.

Baroody, Eiland, and Thompson (2009) conducted an intervention study where preschoolers were instructed for ten weeks, three times a week, in small groups using manipulatives and games that focus on basic number concepts, verbal counting, object counting, and numerical relations. In a second phase, children were randomly assigned to one of three groups for another ten weeks: semi structured discovery learning; structured learning and explicit instruction; and haphazard practice. All groups made significant gains on an early math assessment, although the absence of a non-intervention comparison group makes it difficult to determine whether the gains were due to normal development or to the interventions. There were no reliable group differences in specific skill areas (e.g., mental arithmetic, $n + 1/1 + n$), however.

Despite the importance of early number development and the evidence that it can be taught to young children, the National Research Council's Committee on Early Childhood Mathematics (2009) concludes, "most early childhood programs spend little focused time on mathematics, and most of it is low in instructional quality. Many opportunities are missed for learning mathematics over the course of the preschool day" (p. 339). Thus, many children enter kindergarten with relatively few number experiences. As noted earlier, kindergarten is a key period in a child's education; kindergarten failure can lead to chronic underachievement in math.

Although many kindergarten curricula are now incorporating more mathematics into the school day (Chard et al., 2008), research has shown that high-risk children benefit most from intensive (at least 30 minutes per session) instruction in small-groups of three to six children (Gersten et al., 2007). To address this issue, Dyson, Jordan, & Glutting (2011) developed a small group number sense intervention for at-risk kindergartners. The lessons targeted key number competencies that underlie mathematics difficulties, that is, number, number relations, and number operations. Consistent representations (primarily chips, black dots and fingers) were used because previous work has shown that young children often focus on perceptual variables in tasks rather than on relevant numerical information in mathematics-

related activities (Rousselle, Palmers, & Noël, 2004). Activities were centered on a number list from one to ten (Ramani & Siegler, 2008) and adaptive strategy use was emphasized on addition and subtraction tasks. In a randomized trial with an initial cohort of children, it was found that the intervention group improved in number sense relative to business-as-usual controls, with moderate to large effect sizes; however, improvements were stronger for a numeracy measure more closely aligned to the instruction than for a more general math achievement measure -- which were either not significant at all or not sustained over time (Dyson et al., 2011).

The Present Study

In the present investigation, we report the findings of a randomized controlled trial with a second cohort of kindergartners. The study built on previous work in several ways. Although the number topics were similar to Dyson et al. (2011), we aligned our curriculum more closely to the kindergarten CCSS (Common Core State Standards Initiative, 2010) to provide children with more exposure to skills that they will increasingly be expected to know at first-grade entry (see Table 3). We also adapted instruction to introduce certain activities earlier in the curriculum to allow for greater growth. These activities addressed base-ten concepts, identification of numerals above 20 and their associated quantities, and the concept of before and after a number. Children were exposed to written calculations in more varied situations and contexts (e.g., horizontal and vertical). We expected that these revisions would lead to further gains on more distal math achievement measures as well as on proximal numeracy outcomes.

Second, in addition to a control group of children who received business as usual, we included a small-group control, which provided vocabulary instruction. It is possible that general features of small-group instruction contributed to gains on the assessments in previous work (Dyson et al., 2011) rather than the number activities themselves. Moreover, vocabulary instruction is potentially relevant to low-income learners who experience particular weaknesses with language-based story problems (Jordan, Huttenlocher, & Levine, 1994). There is a strong association between language development and math learning (Cirino, 2011), and math difficulties frequently co-occur with reading and language deficits (Barbarese, Katusic, Colligan, Weaver, & Jacobsen, 2005). Young children construct meanings for numerical symbols through language. For example, knowledge of number words serves as a placeholder for children to build their understanding of natural numbers (Carey & Sarnecka, 2006).

The present study used a pretest, post test, and delayed post test design. Dependent variables included an assessment of number sense, one that is sensitive to change in kindergarten (Jordan, Glutting, Ramineni, & Watkins, 2010), as well as a conventional measure of mathematics achievement and an oral language measure. Because all participants were attending schools in low-income communities, we considered the entire population at risk for underachievement. The work builds on previous investigations through the use of two contrast groups -- language and business as usual -- and its alignment with the 2010 CCSS. It also differs from other successful intervention work in that it is not a full curriculum but rather a short intervention that can be used to augment any kindergarten curriculum.

Method

Participants were full-day kindergartners from five elementary schools in the same school district in Delaware. The schools served mainly low-income children, with the mean participation in the free/reduced price lunch program being 93 percent. Children in all of the schools scored significantly below the state average on third-grade state mathematics test. Children in four of the five participating schools were receiving *Math Connects* (McMillan/McGraw-Hill/Glencoe, 2009) as their mathematics curriculum while the fifth received *Math Trailblazers* (Teaching Integrated Mathematics and Science Project, 2008).

Original participants were 132 kindergartners. Within each school and kindergarten class children were randomly assigned to participate in the number sense group (number sense), a language group (language) or a business-as-usual control group (control) ($n = 44$ in each group) such that children from each classroom were assigned equally to the three groups (e.g., if six students from a classroom participated, two were assigned to each of the three groups). The small groups had the same children for each intervention. During the study period, two children (one from the language and one from the number sense group) moved out of the district; two additional children (one from the language and one from the number sense group) refused to come to the intervention about four weeks into the study.

The final sample had 63 girls (49%) and 65 boys (51%). The school identified 62 students as African American (48%), 52 as Hispanic (41%), 13 as Caucasian (10%), and one as bi-racial. Thirty-two (25%) students were identified as English Language Learners (ELL) and were enrolled in kindergarten ELL classrooms. Table 1 provides descriptive information of the sample broken down by group.

Measures

Number sense—The Number Sense Brief (NSB) (Jordan et al., 2010) assesses number knowledge in the following areas: counting; number recognition; number comparisons; nonverbal calculation; story problems; and number combinations. A perfect raw score is 45.¹

The counting subarea (10 points) asked children to enumerate a set of 5 stars and to give the cardinal value, to give the examiner “4” items, to count to 10, 20, and 30 respectively, to count up from 4 and 8, and to count down from 6 and from 14.

On number recognition (8 points), children were asked to name the following numerals: 4, 9, 13, 16, 28, 37, 82, and 124.

Number comparisons (9 points), adapted from Griffin (2004), asked children to indicate the numbers right after and two after 7; the numbers right before 4 and right before 6; the bigger of 5 and 4 as well as 7 and 9; the smaller of 8 and 6 as well as 5 and 7; and the number that is closest to 5 – 6 or 2. Numerals were shown and named for each item.

¹Several more challenging counting items (i.e., higher counts) and number combinations ($7 + 1$; $10 + 1$; $9 - 1$) were added in the present NSB version to provide additional information.

On nonverbal calculation (4 points), children were shown a horizontal array of black dots. The dots were then hidden with a check box lid with an opening on the side. Dots were either slid in or out of the box. Children were shown 4 arrays of dots in a multiple-choice format and asked to “point to the number of dots in the box now.” The calculations were $2+1$, $3+2$; $4+3$; and $3-1$. Because a common error for children is to give either the first or second addend as the sum (Levine, Jordan, & Huttenlocher, 1992), each addend was given as one of the four choices along with the correct sum. The remaining choice was either one more or one less than the sum, whichever was most appropriate. Following the same idea for the subtraction problem, the choices were 1, 3, 2, and 4 (chosen because children often add instead of subtracting).

Story problems (5 points) were read to children. The problems represented the calculations $2 + 1$, $4 + 3$, $3 + 2$, $6 - 4$, and $5 - 2$. An example of an addition story problem is: Jose has 3 cookies. Sarah gives him 2 more cookies. How many cookies does Jose have now? An example of a subtraction problem is: Paul has 5 oranges. Maria takes away 2 of his oranges. How many oranges does Paul have now? No visual representation was given either of quantity or of numerals used in the story.

Finally, number combinations (9 points) were shown to children and read as “How much is m and n ?” or “How much is m take away n ?” The combinations were $2 + 1$; $3 + 2$; $4 + 3$; $2 + 4$; $7 + 1$; $10 + 1$; $7 - 3$; $6 - 4$; $9 - 1$. Children were shown the written combinations as each problem was presented.

On both NSB story problems and number combinations, children were given a pencil and a piece of paper as well as a number list from 1 to 10 and were told they could use these tools, their fingers, or whatever they wanted to solve the problems. An analysis of all errors was conducted on story problems and number combinations to determine if there were false positives for story problems or number combinations because of response biases. The most common response biases were stating the first or second addend for an addition problem or stating the minuend or subtrahend for a subtraction problem. In no cases did this result in a correct response. No student consistently answered with one greater than the largest addend across all items, which could have resulted in a false positive on $N + 1$ items (i.e., $2 + 1$ or $10 + 1$).

The NSB is internally consistent, with a coefficient alpha of at least .90 over the three study points. Reliabilities by subarea and time are presented in Table 2. (Nonverbal calculation in general and story problems at pretest have low internal consistency, and thus analyses with these subareas should be interpreted cautiously.) The NSB has strong predictive validity in that it is strongly correlated with mathematics achievement and growth in first through third grades (Jordan et al., 2009; Jordan & Glutting, 2012).

Mathematics achievement—Students’ general mathematics skill was measured with the *Woodcock-Johnson III Tests of Achievement* (WJ) Form C Brief Battery: Applied Problems and Calculation subtests (Woodcock, McGrew, Schrank, & Mather, 2007). The WJ Brief Math, a standardized measure with national norms, is highly correlated with the longer

broad math version (.99) in the kindergarten age range. Internal reliability is above .90 for both subtests with kindergartners.

Vocabulary—We selected 15 items (12 intervention words and three everyday words) from four subtests of the Bracken Basic Concept Scale-3: Revised (Bracken, 2006): Sizes/Comparisons, Direction/Position, Quantity, and Time/Sequence. All of these subareas are strongly correlated with receptive vocabulary more generally (Bracken, 2006). A word was read to the child and then s/he was asked to point to the picture that showed the word. Overall internal reliability of the resulting questions was .58 at pretest, .64 at post test, and .73 at delayed post test. (Given the modest estimate at pretest, results should be interpreted cautiously.)

Procedures

Students were tested individually three times, using a pretest, post test, and delayed post test design. Each testing period was completed within two weeks. All measures were administered at each time-point. Pretest measures were given in the late fall of kindergarten. The math and language interventions were conducted over eight weeks between January and March. The 30-minute lessons were carried out in small groups of four, three days per week (Monday, Wednesday, Friday) for a total of 24 lessons. The interventions were typically provided during free center time and did not conflict with regular mathematics or literacy instruction. Students were administered the post tests during the week following the last intervention and again eight weeks later.

Six university graduate students in education who were fully trained in all testing protocols administered research measures. The same research team along with five trained university undergraduate students taught the carefully scripted number sense and language intervention. At post testing and delayed post testing, intervention instructors assessed only students who were *not* part of their intervention groups and were blind to group membership.

To assess whether the intervention procedures were carried out as designed, three intervention lessons were audio-recorded for each of the instructors. The recorded lessons occurred in the early, middle, and late stages of the program. Three graduate research assistants checked the intervention recordings against the written lesson scripts and no substantive deviations were found. During the lesson time, instructors also took notes and logged completion of activities on copies of the lesson script. These scripts were reviewed for completion of tasks. All instructors in both the language and number sense interventions completed all 24 lessons. As a result of the analyses of the particular lessons, it was found that all activities were recorded as completed.

With respect to the individual students, the average attendance was 21 ($SD = 2.8$) lessons for the language group and 22 ($SD = 1.6$) lessons for the number sense intervention group. Specific attendance information was not obtained for the control group, although classroom teachers reported that all of the children attended school regularly during the study period.

Number sense intervention—All intervention instructors were trained on the appropriate instructional approaches and rationale for the lesson activities. Methods highlighted at weekly instructor meetings included how to correct errors, make correct use of lesson materials, incorporate gestures, and maintain student engagement and attention. In order to practice instruction techniques as well as to note any unclear or confusing activities, groups of instructors then implemented intervention lesson activities with each other. Instructors' issues were discussed and resolved with the whole team.

As stated earlier, the 24 lessons were developed and tested in an initial cohort and aligned with the kindergarten CCSS (see Table 3). The lessons used carefully chosen vocabulary (e.g. before, after, plus, minus, bigger, smaller, more, less, altogether) and a compare and contrast approach overall (i.e., before and after, addition and subtraction, $n+1$ and $n-1$ are presented simultaneously). Frequent progress monitoring, through regular, short assessments during the lesson, allowed for prescriptive instruction at regularly planned times during the lessons. For example, a student who has achieved skill proficiency can complete independent counting activities while the instructor works with a child who cannot recognize the numbers to 10.

The lessons drew from the most effective activities in the initial iteration of the study. Although the number sense topics remained the same, some activities (e.g., using Unifix cubes to teach base ten concepts and the introduction of numbers above 20) were introduced earlier to allow for greater growth. More attention was devoted to before/after skills (e.g., number list activities were added that built on children's' experience with "lining up"), which proved to be difficult students in the first cohort. Rote counting activities to the number 20 were reduced because children easily reached the ceiling on this skill in the first cohort and it was determined that children perform this sort of counting regularly in their general classroom instruction. The following are the types of activities implemented:

(1) Number recognition and base ten principles: Each session began with a new, "magic" number. We started with the numeral 10 since most children were proficient in recognizing one-digit numerals. Children were introduced to the base ten structure of number through the use of Unifix cubes (to emphasize quantity) and decade and units cards (to emphasize the symbolic nature of the base ten system). In the first activity, children counted out 10 Unifix cubes and put them into a stick of ten. The children were also introduced to the numeral 10. For subsequent lessons, the next number in the count sequence was introduced as the new number for the day.

To emphasize the structure of the base ten numeration system, the lessons began with counting out the correct number of Unifix cubes and grouping ten cubes into a stick of ten. For example, for 15, the children counted out 15 Unifix cubes, put ten cubes into a stick and the remaining five were left on the table. The instructor circled the stick of ten and said, "See, 15 is ten and five more (circling the pile of five cubes)." The instructor also constructed the numeral using the '10' decade card and single digit overlays (Fuson et al., 2001). For 15, the instructor placed the 5 card over the 0 in 10 and said, "See, 15 is 10 and 5 more." A direct connection was made between the stick of ten and the 10 decade card, as well as between the pile of five blocks and the unit card "5." Similar activities were

introduced in subsequent lessons up to 30. After 30, the children were taught the remaining decades to 90 and were directed to build various two-digit numbers using the decade and single digit cards as well as the Unifix cubes. Since the English naming system for numbers 11 – 19 is inconsistent and difficult to memorize, extra time was given to review these numbers.

Children also practiced number recognition, taking turns, for a short time each day. Each day, the new (“magic”) number was added to the pile. In order to focus attention, children were directed to watch for the “magic” number. When the new number came up, every child was permitted to call it out, whether or not it was his or her turn. Those who responded correctly received a special sticker. If any of the children did not respond correctly, the new number card was returned to the pile until all had success.

(2) Number sequencing: Children used the number recognition cards to participate in number sequencing activities. For example, children were randomly dealt two number recognition cards. Beginning with a number put down by the instructor, children were invited to put down the next number in the sequence if they had the card in their hand. Students took turns putting the cards in sequence, sometimes forward and sometimes backward. The words “before” and “after” were used during these sequencing activities to help strengthen these difficult concepts.

(3) Verbal subitizing: Using cards with various numbers of circles to 4, children were taught to recognize and name quantities instantly without counting. Beginning with two circles, the number of circles on the cards increased as we moved on to larger numbers in the partners’ activities (see Part-Whole relationships). When all children reached proficiency with four circles, the subitizing cards were practiced only periodically as review. Children also practiced making finger quantities to 10 quickly (without counting).

(4) Finger use: Children were taught to count on their fingers but were encouraged early to make the numbers 1–5 with their fingers automatically, without counting. Next, children were shown how to solve both addition and subtraction problems with their fingers. For example, if the combination was $3 + 2$, they practiced showing 3 on one hand and 2 on the other and then counting all. For subtraction, children produced the minuend and folded down fingers for the subtrahend (“take away” model).

Finger counting up from a given start number was practiced and served as an introduction to the “counting-on” strategy. We restricted counting-on problems (both story problems and number combinations) to plus one and plus two since young children seem able to keep track of adding one or two more, but adding more than two requires a more sophisticated system of keeping track (Fuson & Secada, 1986). Children were encouraged to automatically make the first addend on their fingers and then count up one or two more while putting up one or two more fingers. We also encouraged counting-on when making numbers greater than five on fingers. Children were reminded that they did not need to count the fingers one hand and could start counting with five when showing numbers greater than five.

(5) Number list activities: Activities using a number list from 1 to 10 supported the learning of a) counting principles related to one-to-correspondence, stable order, and cardinality, b) right before and right after, c) the $n + 1$ principle, as well as $n - 1$, and d) numerical magnitude comparisons. A number list with accumulated quantities above each number, known as a *cardinality chart*, was also used as a visual aid for magnitude comparison (bigger/smaller) activities.

(6) Written number activities: Each number list activity was repeated using only symbolic representations of number. For example, a child would be shown a card with two numerals on it and be asked, “Which is bigger?” or, a group might be shown a card with the numeral five and be asked, “Who is holding the number that comes right before five?”

(7) Part-whole relationships: To emphasize that numbers are made up of smaller sets, partner cards showing the part-whole relationships for the numbers 1–5 were used (e.g., for the number 4, one partner card showed a row of four dots separated into two sets of two dots and another showed a row of four dots separated into a set of three dots and one dot). Partner card activities included connecting partners to written combinations with Arabic numerals, both for addition and subtraction. By rotating the partner card 180 degrees, the instructor demonstrated the commutative property of addition.

(8) Problem solving and operations: Children were encouraged to use adaptive strategies to solve simple story problems and number combinations. As stated above, children were encouraged to use their fingers as one way to solve addition and subtraction problems. Combinations went up to sums totaling 10 (and differences with a maximum minuend of 10) so that children could perform operations using their fingers. Children were also shown how to draw a model to solve problems and how to use the number list for counting on or counting down. In order to support connections between concrete and symbolic representations of number, children were asked to connect story problems to a written combination as well as the appropriate partner card. Children were exposed to conventional vertical and horizontal written formats.

(10) Linear number board game: At the end of each session, children were rewarded for their hard work by playing the *Great Race Game*, adapted from Ramani and Siegler (2008). The game uses a colorful number list from 1 to 10. Children spun a spinner to determine if they moved one or two spaces forward or backward. Instructors were careful to use the words “before/after” and “more/less” during the game to help the children connect those vocabulary words to movement on the number list. Each child was given an individual small game board to use as well as their own token. This allowed everyone to see the board adequately, to focus on counting, and to avert behavioral issues related to grabbing tokens on a shared game board, etc.

Language intervention—The language intervention was parallel to the number sense intervention, only it focused on storybooks with some quantitative vocabulary. Children were exposed to 43 vocabulary words from eight different storybooks. Quantitative words such as *before/after*, *big/small*, and *part/whole* were taught alongside general vocabulary words such as *wring*, *greedy*, *probably*, *discover*, and *alone*. The activities in the language

intervention were adapted from Beck and McKeown's (2001b) *Text Talk: Level A* curriculum. This curriculum was developed specifically for children from low SES backgrounds (Beck & McKeown, 2001a). Children's storybooks are a natural resource for vocabulary learning (Beck, McKeown, & Kucan, 2002), generating many word-based activities related to the content of the story.

Business as usual control condition—The business-as-usual control groups received their regular math instruction each day along with the intervention groups. During the intervention time, business as usual controls were engaged in regular kindergarten activities, typically center work or special subjects.

Data Analysis

A series of 22 multilevel models (MLMs) were employed, one for each dependent variable. The dependent variables, at 2 time points, were the NSB brief total, NSB subareas (counting, number recognition, number comparisons, nonverbal calculation, story problems, number combinations), vocabulary, WJ total, WJ Applied Problems, and WJ Calculation. Data were analyzed using the Proc Mixed procedure from SAS (Littel, Milliken, Stroup, Wolfinger, & Schabenberger, 2006). Children within classrooms were randomly assigned to treatments (number sense, language, and control). Treatments were then delivered to small groups defined by children's classroom assignments. Children remained in the same cohort throughout the intervention. Therefore, data were analyzed using two-level MLMs. Level one was comprised by individual student variables. Level two contained the nesting variable of cohort/classroom.

Treatment was a level one variable and not a level two variable because the study did not employ a clustered, random-assignment design. The three treatment groups were dummy coded into two predictors (number sense vs. controls and language vs. controls). Other level one predictors served as covariates and included: ELL (yes, no), gender (with females serving as the base group), and pretest scores from both the NSB and the WJ total. Interval-level predictors were grand-mean centered (Hox, 2002; Kreft & de Leeuw, 1998).

The co-varied pretest scores from the NSB and/or pretest scores from WJ total served to minimize potential confounding that might be attributable to prior mathematics knowledge among the groups. The pretest covariates also reduced unexplained variance, thereby increasing the power of the analyses to detect treatment effects (Field, 2009; Maxwell & Delaney, 2004). Three combinations of pretest covariates were used: (a) pretest scores from the NSB, (b) pretest scores from the WJ total, and (c) pretest scores from both the NSB and WJ total. The reason for including a variety and combination of covariates was due to concerns about multicollinearity. However, the interpretations remained the same. Therefore, results are presented here for only those models where pretest scores from both NSB and WJ total were included.²

The initial model for each achievement area was unconditional. This model is also sometimes referred to as a "random intercept model" that predicts the level one intercept of

²Results with the NSB and WJ pretest scores as separate covariates can be obtained from the authors.

the dependent variable as a random effect of the level two grouping variable (Garson, in press, p. 8). More specifically, the initial model entered only the nesting variable at level two (i.e., the three treatment groups). The unconditional analyses included those assuming completely generalized (unstructured) error components (Wolfinger, 1993, 1996). However, convergence problems took place during eight of the 22 analyses. A simpler (diagonal) covariance structure was then specified. The net effect is that the final models employed what Raudenbush and Bryk (2002) refer to as “one-way ANCOVA[s] with random effects” (p. 25) where the random slope effects were fixed.

Subsequent conditional models sequentially added and tested covariates under full maximum-likelihood constraints, and the final model for each achievement area featured only statistically significant random and fixed effects that were extracted through restricted maximum-likelihood estimation (per Littell et al., 2006; Milliken & Johnson, 2002). The sequencing of covariate entry proceeded as recommended by Bauer and Curran (2006).

Post hoc comparisons of outcomes were based on contrasts of least-squares means (means adjusted for any group imbalance). Where statistically significant fixed effects were discovered, effect sizes (Cohen’s *d*) were estimated using least-squares means and standard deviations. Least-squares means (Searle, Speed, & Milliken, 1980) are population parameter estimates and do not have standard deviations (*SDs*) as are available for ordinary sampling distributions. Effect size estimates, such as Cohen’s *d* and Hedges’ *G*, require an *SD* to specify the measurement scale in the formula denominator (Rosenthal, 1994). Typically, *SDs* specific to each given contrast group would be pooled to estimate the proper *SD* for *d* or *G*. Such a correction would be inappropriate here because the least-squares means are by their nature already corrected for contrast group imbalance, such that pooled *SDs* would no longer represent the measurement scale and would tend to bias estimates. Alternatively, the *SD* for the predicted scores of all children was applied here in accordance with the method developed by McDermott (2011) and applied previously by Fantuzzo, Gadsden and McDermott (2010).

Corrections for multiple comparisons followed the guidelines developed for the Institute of Education Sciences (IES) (Schochet, 2008). Specifically, analyses directed to overall scores from instruments were considered to be primary (“confirmatory” in IES terminology, p. 4), whereas analyses completed with subscores were considered to be exploratory. The following six dependent variables were considered primary: NSB Total post, NSB Total delayed, WJ Total post, WJ Total delayed, Total Vocabulary post, and Total Vocabulary delayed. The confirmatory analysis addressed the multiple comparison issue by employing Holm’s (1979) correction where the overall number of comparisons (*k*) equaled seven. No correction was applied to the exploratory analyses, as per IES recommendations.

Results

Table 4 provides raw score means (*Ms*) and *SDs* for the dependent variables (NSB, WJ, and Vocabulary) by group. Likewise, the *Ms* and *SDs* are separated by pretest, post test, and delayed post test. Preliminary comparisons were conducted using analyses of covariance and revealed no statistically significant effects involving language status (ELL vs. non-ELL),

gender, or attendance (number sense vs. language). Therefore, the variables were excluded in subsequent analyses. The deletions, in turn, made the models more parsimonious and easier to interpret.

Intraclass correlation coefficients (ρ s) were calculated for the unconditional MLMs. Results across the 22 analyses showed an $M \rho = .11$ and $SD = .14$. Results, on average, indicate that 11% of the variance in the achievement outcomes existed between nesting variable of cohort/classroom. In addition, 12 of the coefficients exceeded a value of .05 that is often used as the cut point for determining the need for an MLM (Raudenbush, Martinez, & Spybrook, 2007; Snijders, 2005).

Table 5 provides p values and effect sizes (i.e., Cohen's [1988] d statistic) for the MLMs. In all, 22 statistical comparisons are presented in Table 5. Interestingly, regardless of whether a specific analysis was statistically significant, none of the comparisons revealed even a single instance where the control group outperformed the number sense intervention group on a math measure. Of the 20 comparisons with math measures, 14 (70%) were statistically significant in favor of the number sense group. More importantly, all four of the confirmatory analyses were statistically significant in favor of the number sense group. (In contrast, the number sense group performed significantly worse than the control group on the confirmatory vocabulary measure at post test and at the delayed post test.)

An analysis was completed of the obtained effect sizes for the 14 analyses where statistically significant math results were obtained. The $M d$ was 1.12. This exceeds Cohen's (1988) benchmark for a large effect size. Of the 14 statistically significant analyses, none showed a small effect size; six (43%) showed a medium effect size; and eight (57%) showed a large effect size. The overall effect size was repeated for the same 14 analyses, but it was calculated between the Language intervention group versus the controls. Here, the $M d$ was -0.18 . This averaged effect size indicated that the language intervention was *less effective* in raising outcome scores than the control group. Lastly, it is possible to compare the two averaged effect sizes. Results show a very large effect: $d = 1.3$, i.e., $1.12 - (-0.18) = 1.3$. Therefore, it is reasonable to infer that the number sense intervention was much more effective in raising outcome scores than the language intervention.

It is noteworthy that the number sense intervention group's advantage remained significant at delayed post test, although reduced. The number sense intervention children made particularly strong and sustainable gains on story problems, with large effects both at post (2.64) and delayed post test (2.27). There were no significant effects on counting or on nonverbal calculation. On the WJ Applied subtest (which includes a broader range of story problems), no significant effects were seen at immediate post test but there were significant and medium-sized ones favoring the number sense group at delayed post test. The number sense group also showed significant and meaningful gains in number combinations and number knowledge, relative to controls. These gains are further reflected in large to medium performance gains on the WJ Calculation subtest.

To put the findings into context, the mean WJ math percentile scores (based on national age norms) by group and time of testing are presented in Table 6. In practical terms, the number

sense intervention group, on average, went from the 8th percentile in WJ calculation at pretest to the 45th percentile at delayed post test; the change was the 18th to 31st percentile for the language group and the 18th to 37th percentile for the control. The relatively large percentile gains made by all groups in WJ calculation seem to reflect floors at pretest (i.e., all groups averaged a pretest raw score of 1 or less; see Table 4). On WJ applied problems, the number sense group went from the 33rd to the 42nd percentile, the language group from the 31st to 33rd percentile, and the control group from the 35th to 36th percentile.

Discussion

Math achievement in low-income children is mediated by their performance and growth in number sense during kindergarten (Jordan et al., 2009). Early number experiences provide children with the knowledge they need for success in formal mathematics (National Research Council, 2009). The aim of the present study was to improve kindergartners' whole number competencies as reflected by performance on a numeracy measure that is sensitive to growth in kindergarten as well as on more general achievement indicators. Controlling for initial skill level, children who received the number sense intervention performed better than controls, with statistically meaningful effects on most measures. Many of the effects held eight weeks after the intervention was completed, suggesting that children internalized what they learned. Given the importance of kindergarten performance to long-term outcomes (e.g. Duncan et al., 2007), the findings are especially important.

Children started the eight-week, 24 session intervention in mid-kindergarten. At pretest, most children had foundational knowledge for building number sense and connecting ideas. They could count to ten but were not as successful with higher counts. Nearly all recognized one-digit numbers but not ones with two or more digits. Most children could name the number after a digit and which of two digits is smaller or bigger, although many had trouble naming the number before a digit and the number two after a digit. While many children calculated successfully in a nonverbal format with object representations, few showed facility with arithmetic story problems and number combinations.

Although children in all groups made gains in number skills during the study period, the number sense group clearly outperformed the controls in most areas. Particularly noteworthy were the consistently large effect sizes on story problems ($d = .2.64$ at post test and 2.27 at delayed post test when controlling for achievement at pretest), where the number sense group gained more than the controls. At delayed post test, for example, 81% of children in the number sense group could solve the change plus story problem, "Jose has 3 cookies. Sarah gives him 2 more cookies. How many cookies does Jose have now?" and 74% could solve the change minus problem, "Paul has 5 oranges. Maria takes away 2 of his oranges. How many oranges does Paul have now?" versus 45% and 36% for the language group and 43% and 36% for the control group. Without any intervention, an earlier study showed that low-income kindergartners are 4 times more likely to fall into a low performance, flat growth group on story problems than are middle-income kindergartners (Jordan et al., 2006). We compared the number sense children's performance on story problems in the present study with that of middle-income children in a former study (i.e., Jordan et al., 2010) on a set of items that were common to both studies. In the spring of kindergarten middle-income

children on average were able to solve about 40% of the story problems vs. 80% for our intervention children at delayed post test.

The intervention group's gains on number combinations also are noteworthy, especially in light of the significant and unique relationship between knowledge of number combinations in kindergarten and the development of later number fact fluency (Locuniak & Jordan, 2008). Kindergarten knowledge of addition number combinations is a marker for mathematics learning disabilities (Mazzocco & Thompson, 2005). A closer look at children's performance on individual problems at delayed post test reveals that the intervention group performed better, on average, than both the language and control groups on all items. The majority of the number sense children were able to add 1 to a larger number (e.g., $10 + 1$) at post test, suggesting they can construct the number-after rule for adding one (10 and 1 is the number after 10, that is, 11) (Baroody et al., 2009). Subtraction problems were hardest for all children, regardless of group membership.

There were no significant group effects on nonverbal calculation and counting and more modest effects on number recognition. All children could solve most of the nonverbal items at delayed post, suggesting that a ceiling effect made the task less sensitive for kindergarten. [It is important to note that the nonverbal measure was generally less reliable than the other areas, so findings should be interpreted cautiously.] With respect to counting, all children gained at about the same rate, most likely because counting activities are emphasized in typical kindergarten math curricula. By delayed post test, most children could count to 30; however, we did not assess children's counting to 100 or by 10s, a kindergarten goal of the CCSS (Common Core State Standards Initiative, 2010). Counting backwards from 14 was hard for nearly everyone regardless of group membership. Although counting down from a number is a potential strategy for solving subtraction problems, it is developmentally difficult for many young children and often results in errors (Baroody, 1984; Fuson, 1986). In the present intervention, we emphasized inverse operations and encouraged children to solve plus one and plus two addition problems using a fingers counting-on approach, a precursor to solving subtraction problems by counting up (Fuson & Willis, 1988). Although counting on using fingers is amenable to instruction (Fuson & Secada, 1986), further research should examine whether children can make the transition to counting on mentally, without the aid of fingers.

With respect to number recognition (where the number sense children outperformed the other children right after the intervention but not two months later), all groups gained substantially in their ability to recognize two-digit numerals (e.g., 7% of number sense children could name the numeral 82 at pretest and 79% at delayed post, 10% and 69% for language children, 11% to 73% for controls) while most were still inaccurate with three-digit numerals (e.g., about a third of the children in each group could name 124) at the end of kindergarten. These results suggest recognition of two-digit numbers, a conventional skill, receives attention in kindergarten more generally.

Achievement effects favoring the number sense children on the more general WJ math achievement test (Woodcock et al., 2007) were encouraging and suggest skill carryover, a substantial improvement over previous work where transfer was more limited (Dyson et al.,

2011). Controlling for initial number sense and general math skills, the intervention children outperformed the control group on total achievement at both post tests. Intervention effects also were observed on applied problems and calculation subtests. Based on WJ (Woodcock et al., 2007) national kindergarten norms, children started out especially low in Calculation, with a mean percentile score of 15 for the whole group (vs. 33rd percentile on applied problems). Although (by chance) the number sense intervention group, on average, started about 10 percentile points lower than the controls, they finished about 10 points higher. Early WJ Calculation items are strictly conventional with written addition and subtraction items in both horizontal and vertical formats. Early WJ Applied Problems items require children to count different types of items in pictures and to solve simple word problems in various contexts. It is likely that the number sense children's increased knowledge of addition/subtraction operations (as seen on the NSB) is reflected in their improvements the WJ math achievement measure.

We included a language condition that did not teach number concepts directly to examine the impact of small group instruction more generally. It could be argued, for example, that the special attention of working in small groups helps young children build general skills, such as task focus or attention to instructions, which would improve performance on number assessments. This was not the case in the present study, as the language children did not differ from controls on any math outcomes. Neither the number sense nor the language group differed from controls on the vocabulary measure at either post test; thus it is not clear from the present study whether language gains derived from special instruction lead indirectly to gains in number sense. Future studies should use more varied and reliable measures of language (general vocabulary as well as quantitative vocabulary) and perhaps provide language instruction that is targeted directly to quantitative concepts.

Overall, our findings can be interpreted relative to the developmental steps in the number, relations, and operations cores described by the National Research Council (2009). Most kindergartners in our study showed competence at the end of the year in the number core steps (i.e., subitizing, set enumeration, cardinality, and saying and counting two digits) and many in the relations core (i.e., matching or counting to find out which is more and which is less for two numbers – 10). Where the intervention really made a difference was in the addition/subtraction operations core related to using numerical strategies to solve story, oral number word, and written numeral problems with totals to 10. Children with these competencies should be well equipped to operate on numbers with totals to 18 in first grade and for thinking about more complex problems or situations, such as finding an unknown addend (e.g., $6 + ? = 8$).

A potential limitation of the study is that the business as usual control group did not receive an additional dose of regular curriculum math to control for instructional time. It is possible that adding thirty minutes of business as usual math instruction during the study period would have led to more gains for the control children. It would also be interesting to see if our number sense curriculum could be adapted as a supplement for large groups to serve more children. In any case, our study reveals that high-risk children can make notable gains in math after only eight weeks of intervention.

Our eight-week intervention model has strong potential for use in response to intervention (RTI) programs for identifying students with genuine learning disabilities versus difficulties related to inadequate instructional support (Fletcher & Vaughn, 2009; National Center on Response of Intervention (NCRTI), 2010). The intervention is aligned with the CCSS (Common Core State Standards Initiative, 2010) so should support most kindergarten math curricula. Children who continue to have weaknesses in the number and relations core (as observed on the NSB or a similar measure) after receiving the intervention are most at risk for learning disabilities and may warrant referral for a special education evaluation. Children who make adequate progress in the number and relations core but not in operations may need additional help beyond the eight-week kindergarten period. Even for at-risk children who make good progress in all core areas, it may be unrealistic to expect similar levels of growth in subsequent grades without targeted follow-through (Clements & Sarama, 2009).

Acknowledgments

This research was supported by Grant # HD059170 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. We are grateful to the participating children and teachers for their generous cooperation.

References

- Abedi J. The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*. 2004; 33:4–14.
- Barbaresi, MJ.; Katusic, SK.; Colligan, RC.; Weaver, AL.; Jacobsen, SJ. Math learning disorder: Incidence in a population-based birth cohort, 1976–1982. Vol. 5. Rochester: Minn; 2005.
- Baroody AJ. An investigative approach to the mathematics instruction of children classified as learning disabled. *Educational context of learning disabilities*. 1984; Chapter 14:545–615.
- Baroody AJ. The development of counting strategies for single-digit addition. *Journal for Research in Mathematics Education*. 1987; 18(2):141–157.
- Baroody AJ, Eiland M, Thompson B. Fostering at-risk preschoolers' number sense. *Early Education and Development*. 2009; 20:80–128.
- Bauer, DJ.; Curran, PJ. Multilevel modeling of hierarchical and longitudinal data using SAS: Course notes. Cary, NC: SAS Institute; 2006.
- Beck IL, McKeown MG. Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher*. 2001a; 55(1):10–20.
- Beck, IL.; McKeown, MG. Text Talk: Robust Vocabulary Instruction: Level A. New York: Scholastic, Inc; 2001b.
- Beck, IL.; McKeown, MG.; Kucan, L. Bringing words to life: Robust vocabulary instruction. New York: The Guilford Press; 2002.
- Bracken, BA. Bracken Basic Concept Scale-Third Edition: Receptive. San Antonio, TX: The Psychological Corporation; 2006.
- Butterworth, B.; Reigosa, V. Information processing deficits in dyscalculia. In: Berch, DB.; Mazzocco, MMM., editors. Why is math so hard for some children?. Baltimore, MD: Paul H. Brookes Publishing Co; 2007. p. 65-81.
- Carey, S.; Sarnecka, BW. The development of human conceptual representations: A case study. In: Munakata, Y.; Johnson, MH., editors. Processes of change in brain and cognitive development: Attention and performance. Vol. XXI. Oxford, UK: Oxford University Press; 2006. p. 473-496.
- Case, R.; Griffin, S. Child cognitive development: The role of central conceptual structures in the development of scientific and social thoughts. In: Hauert, CA., editor. *Advances in psychology-Developmental psychology: Cognitive, perception-motor, and neurological perspectives*. Amsterdam: North Holland; 1990.

- Chard DJ, Baker SK, Clarke B, Jungjohann K, Davis K, Smolkowski K. Preventing Early Mathematics Difficulties: The Feasibility of a Rigorous Kindergarten Mathematics Curriculum. *Learning Disability Quarterly*. 2008; 31(1):11–20.
- Chetty R, Friedman JN, Hilger N, Saez E, Whitmore D, Schanzenbach &, Yagan D. \$320,000 Kindergarten Teachers. *Phi Delta Kappan*. 2010; 92(3):22–25.
- Cirino PT. The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology*. 2011; 108:713–733. [PubMed: 21194711]
- Clarke B, Shinn MR. A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*. 2004; 33(2):234–248.
- Clements DH, Sarama J. Effects of a preschool mathematics curriculum: summative research on the Building Blocks project. *Journal for Research in Mathematics Education*. 2007; 38:136–163.
- Clements DH, Sarama J. Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Education Research Journal*. 2008; 45(2):443–494.
- Clements, DH.; Sarama, J. *Learning and teaching early math: The learning trajectories approach*. New York: Routledge; 2009.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1988.
- Common Core State Standards Initiative. *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers; 2010.
- Dobbs J, Doctoroff GL, Fisher PH, Arnold DH. The association between preschool children's socio-emotional functioning and their mathematic skills. *Applied Developmental Psychology*. 2006; 27:97–108.
- Duncan, GJ.; Brooks-Gunn, J. Poverty, welfare reform, and children's achievement. In: Biddle, B., editor. *Social class, poverty and education: Policy and practice*. New York, NY: Routledge Almer; 2001. p. 49-76.
- Duncan GJ, Dowsett CJ, Classens A, Magnuson K, Huston AC, Klebanov P, Japel C. School readiness and later achievement. *Developmental Psychology*. 2007; 43(6):1428–1446. [PubMed: 18020822]
- Dyson NI, Jordan NC, Glutting J. A Number Sense Intervention for Kindergartners At Risk for Math Difficulties. *Journal of Learning Disabilities*. 2011 Advance online publication. 10.1177/0022219411410233
- Fantuzzo JW, Gadsden VL, McDermott PA. An integrated curriculum to improve mathematics, language, and literacy for head start children. *American Education Research Journal*. 2010; 48 (3): 763–793.
- Field, A. *Discovering statistics using SPSS*. 3. London: Sage; 2009.
- Fletcher JM, Vaughn S. Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives*. 2009; 3:30–37. [PubMed: 21765862]
- Fuson KC. Teaching children to subtract by counting up. *Journal for Research in Mathematics Education*. 1986; 17:172–189.
- Fuson KC, Grandau L, Sugiyama PA. Achievable Numerical Understandings for All Young Children. *Teaching Children Mathematics*. 2001; 7(9):522–526.
- Fuson KC, Secada WG. Teaching children to add by counting with finger patterns. *Cognition and Instruction*. 1986; 3:229–260.
- Fuson KC, Willis GB. Subtracting by Counting Up: More Evidence. *Journal for Research in Mathematics Education*. 1988; 19(5):402–420.
- Garson, GD. *Hierarchical linearling modeling: Guide and applications*. Los Angeles: Sage; in press
- Gelman, R.; Gallistel, CR. *The child's understanding of number*. Cambridge, MA: Harvard University Press; 1978.
- Gersten, R.; Baker, SK.; Shanahan, T.; Linan-Thompson, S.; Collins, P.; Scarcella, R. *Effective Literacy and English Language Instruction for English Learners in the Elementary Grades: A Practice Guide (NCEE 2007-4011)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education; 2007.

- Ginsburg HP, Russell RL. Social class and racial influences on early mathematical thinking. *Monographs of the Society for Research in Child Development*. 1981; 46(6):1–69. Serial No. 193.
- Griffin S. Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly*. 2004; 19:173–180.
- Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979; 6:65–70.
- Hox, J. *Multilevel analysis: Techniques and applications*. Mahwah, New Jersey: Lawrence Erlbaum; 2002.
- Jordan, NC.; Glutting, J. *Number Sense Screener User's Guide*. Baltimore: Paul H. Brookes; 2012.
- Jordan NC, Glutting J, Ramineni C, Watkins MW. Validating a number sense screening tool for use in kindergarten and first grade: Prediction of mathematics proficiency in third grade. *School Psychology Review*. 2010; 39:181–195.
- Jordan NC, Huttenlocher J, Levine SC. Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology*. 1992; 28(4):644–653.
- Jordan NC, Huttenlocher J, Levine SC. Assessing early arithmetic abilities: Effects of verbal and nonverbal response types on the calculation performance of middle- and low-income children. *Learning and Individual Differences*. 1994; 6(4):413–432.
- Jordan NC, Kaplan D, Locuniak MN, Ramineni C. Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*. 2007; 22(1): 36–46.
- Jordan NC, Kaplan D, Oláh L, Locuniak MN. Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*. 2006; 77:153–175. [PubMed: 16460531]
- Jordan NC, Kaplan D, Ramineni C, Locuniak MN. Development of number combination skill in the early school years: When do fingers help? *Developmental Science*. 2008; 11(5):662–668. [PubMed: 18801121]
- Jordan NC, Kaplan D, Ramineni C, Locuniak MN. Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes. *Developmental Psychology*. 2009; 3(45):850–867. [PubMed: 19413436]
- Kempert S, Saalbach H, Hardy I. Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*. 2011; 103(3):547–561.
- Klein, A.; Starkey, P. Fostering preschool children's mathematical knowledge: Findings from the Berkeley math readiness project. In: Clements, DH.; Sarama, J.; DiBase, AM., editors. *Engaging young children in mathematics: Standards for early childhood mathematics education*. Mahwah, NJ: Lawrence Erlbaum Associates; 2008. p. 343-360.
- Klibanoff RS, Levine SC, Huttenlocher J, Vasilyeva M, Hedges LV. Preschool children's mathematical knowledge: The effect of teacher "Math Talk". *Developmental Psychology*. 2006; 42(1):59–69. [PubMed: 16420118]
- Kreft Ita, GG.; Jan de; de Leeuw. *Introducing multi-level modeling*. Thousand Oaks, CA: Sage Publications; 1998. *Introducing Statistical Methods Series*
- Le Corre M, Carey S. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*. 2007; 105:395–438. [PubMed: 17208214]
- Lembke E, Foegen A. Identifying Early Numeracy Indicators for Kindergarten and First-Grade Students. *Learning Disabilities Research & Practice*. 2009; 24(1):12–20.
- Levine SC, Jordan NC, Huttenlocher J. Development of calculation abilities in young children. *Journal of Experimental Child Psychology*. 1992; 53:72–103. [PubMed: 1545190]
- Libertus ME, Feigenson L, Halberda J. Preschool acuity of the approximate number system correlates with school math ability. *Developmental science*. 2011:1–9. [PubMed: 21159082]
- Littell, RC.; Milliken, GA.; Stroup, WW.; Wolfinger, RD.; Schabenberger, O. *SAS systems for mixed models*. 2. Cary, NC: SAS Institute; 2006.
- Locuniak MN, Jordan NC. Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*. 2008; 41(5):451–459. [PubMed: 18768776]

- McDermott, PA. personal communication. Dec 2. 2011
- Maxwell, SE.; Delaney, HD. Designing experiments and analyzing data: A model comparison perspective. 2. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
- Mazzocco MMM, Feigenson L, Halberda J. Impaired acuity of the approximate number system underlies mathematical learning disability. *Child Development*. 2011; 82(4):1224–1237. [PubMed: 21679173]
- Mazzocco MMM, Thompson RE. Kindergarten predictors of math learning disability. *Learning Disabilities Research and Practice*. 2005; 20(3):142–155. [PubMed: 20084182]
- McMillan/McGraw-Hill/Glencoe. Math Connects. Columbus, OH: The McGraw Hill Companies, Inc; 2009.
- Milliken, GA.; Johnson, DE. Analysis of messy data: Vol. 3. Analysis of covariance. New York, NY: Chapman & Hall/CRC; 2002.
- National Center on Response of Intervention (NCRTI). Essential components of RTI—A closer look at response to intervention. Washington, DC: Author; 2010.
- National Mathematics Advisory Panel. Foundations for Success: The Final Report of the National Mathematics Advisory Panel. Washington, D.C: U.S. Department of Education; 2008.
- National Research Council. Mathematics Learning in Early Childhood: Paths toward Excellence and Equity. Washington, D.C: The National Academies Press; 2009.
- Ramani GB, Siegler RS. Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*. 2008; 79:375–394. [PubMed: 18366429]
- Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2. Thousand Oaks, CA: Sage; 2002.
- Raudenbush SW, Martinez A, Spybrook J. Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*. 2007; 29:5–29.
- Rips LJ, Bloomfield A, Asmuth J. From numerical concepts to concepts of number. *Behavioral and Brain Sciences*. 2008; 31:623–642. [PubMed: 19077327]
- Rosenthal, R. Parametric measures of effect size. In: Cooper, H.; Hedges, LV., editors. The handbook of research synthesis. New York, NY: Sage; 1994. p. 231-244.
- Rousselle L, Noël M-P. Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number. 2007
- Rousselle L, Palmers E, Noël M-P. Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*. 2004; 87:57–84. [PubMed: 14698689]
- Royer, JM.; Walles, R. Influences of gender, ethnicity, and motivation on mathematical performance. In: Berch, DB.; Mazzocco, MMM., editors. Why is math so hard for some children?. Baltimore, MD: Paul H. Brookes Publishing Co; 2007. p. 349-367.
- Schaeffer B, Eggleston VH, Scott JL. Number development in young children. *Cognitive Psychology*. 1974; 6:357–379.
- Schochet, PZ. Guidelines for multiple testing in experimental evaluations of educational interventions. Institute of Education Sciences U.S. Department of Education by Mathematica Policy Research, Inc; 2008. Submitted to Retrieved from <http://www.policyarchive.org/handle/10207/8802>
- Searle SR, Speed FM, Milliken GA. Populations marginal means in the linear model: An alternative to least squares means. *American Statistician*. 1980; 34:216–221.
- Siegler, RS.; Jenkins, E. How children discover new strategies. Hillsdale, NJ: Erlbaum; 1989.
- Snijders, TAB. Power and sample size in multilevel linear models. In: Everitt, BS.; Howell, DC., editors. Encyclopedia of statistics in behavioral science. Vol. 3. New York, NY: Wiley; 2005. p. 1570-1573.
- Starkey P, Klein A, Wakeley P. Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*. 2004; 19:99–120.
- Teaching Integrated Mathematics and Science Project. Math Trailblazers. 3. Dubuque, IA: Kendall Hunt; 2008.

- Wolfinger RD. Covariance structure selection in general mixed models. *Communications in Statistics-Simulations*. 1993; 22:1079–1106.
- Wolfinger RD. Heterogeneous variance: Covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics*. 1996; 1:205–230.
- Woodcock, RW.; McGrew, KS.; Schrank, FA.; Mather, N. Woodcock-Johnson III Normative Update. Rolling Meadows, IL: Riverside Publishing; 2007.

Table 1

Demographics by Group

	N	Gender		Ethnicity			Percent ELL	Age (Sept. 1)	
		Male	Female	AA	H	C		Mean	SD
Number Sense	42	25	17	19	17	6	0	28.6	3.59
Language	42	15	27	24	17	1	0	28.6	5.32
Control	44	25	19	19	18	6	1	18.2	4.12
Total	128	65	63	62	52	13	1	25.0	4.38

Note. N = Number of participants, AA= African American, H=Hispanic, C=Caucasian, Other = 1 Bi-racial child, ELL= English-language Learners, SD = Standard Deviation.

Table 2

Reliability of NSB Subtest & Total using Cronbach's Alpha

Subtest	Pretest	Post	Delayed
Counting skills (n=10)	.77	.75	.71
Number recognition (n=8)	.83	.81	.81
Number comparisons (n=9)	.64	.67	.63
Nonverbal calculation (n=4)	.61	.36	.29
Story problems (n=5)	.45	.84	.80
Number combinations (n=9)	.82	.91	.89
Total NSB (n=45)	.90	.92	.91

Note. NSB = Number Sense Brief

Table 3**Number Sense Intervention Alignment with Common Core State Standards**

Common Core State Standards	Number Sense Intervention Activity
Counting and Cardinality	
<i>Know number names and the count sequence.</i>	
1. Count to 100 by ones and by tens	<ul style="list-style-type: none"> Count by ones to 30 Count by ones starting at a given decade Count by tens to 100 using decade cards Recognition of the decades up to 100 Recognition of the numerals 1–100 Sequencing numbers to 30 – forwards and backwards
2. Count forward beginning from a given number within the known sequence (instead of having to begin at 1).	<ul style="list-style-type: none"> Great Race game Building numbers 1–30 using Unifix cubes– count forward from a decade number “Counting on” for solving combinations and story problems
3. Write numbers from 0 to 20. Represent a number of objects with a written numeral 0–20 (with 0 representing a count of no objects).	<ul style="list-style-type: none"> Writing numeral for quantity shown Writing combinations Writing before/after numbers Writing missing numbers in Number List
<i>Count to tell the number of objects.</i>	
4. Understand the relationship between numbers and quantities; connect counting to cardinality. <ul style="list-style-type: none"> When counting objects, say the number names in the standard order, pairing each object with one and only one number name and each number name with one and only one object. Understand that the last number name said tells the number of objects counted. The number of objects is the same regardless of their arrangement or the order in which they were counted. Understand that each successive number name refers to a quantity that is one larger. 	<ul style="list-style-type: none"> Magic Number Activities – counting Unifix cubes Counting fingers, objects Counting objects in a drawing Magic number activities – counting Unifix cubes Count items and write corresponding numeral Adding 1 to n using fingers and number list Language-right <i>before</i> and <i>right after</i> Sequencing numbers 1–20 using <i>after</i> numbers and <i>one more</i> Great Race Game <i>Plus one</i> activities
5. Count to answer “how many?” questions about as many as 20 things arranged in a line, a rectangular array, or a circle, or as many as 10 things in a scattered configuration; given a number from 1– 20, count out that many objects	<ul style="list-style-type: none"> Magic number activities - counting Unifix cubes Counting fingers, objects Counting objects in a drawing
<i>Compare numbers.</i>	

Common Core State Standards	Number Sense Intervention Activity
6. Identify whether the number of objects in one group is greater than, less than, or equal to the number of objects in another group, e.g., by using matching and counting strategies.	<ul style="list-style-type: none"> • Bigger/Smaller using Cardinality Chart • Bigger/Smaller using Number List
7. Compare two numbers between 1 and 10 presented as written numerals.	<ul style="list-style-type: none"> • Bigger/smaller numbers with bigger/smaller numeral cards
Operations and Algebraic Thinking	
1. Represent addition and subtraction with objects, fingers, mental images, drawings, sounds (e.g., claps), acting out situations, verbal explanations, expressions, or equations.	<ul style="list-style-type: none"> • Making drawings to solve combinations • Solving combinations on number list • Solving combinations on fingers
2. Solve addition and subtraction word problems, and add and subtract within 10, e.g., by using objects or drawings to represent the problem.	<ul style="list-style-type: none"> • Story problems: Plus one / Minus one • Making models for story problems • Connecting story problems to combinations • Connecting story problems to partner cards
3. Decompose numbers less than or equal to 10 into pairs in more than one way, e.g., by using objects or drawings, and record each decomposition by a drawing or equation (e.g., $5 = 2 + 3$ and $5 = 4 + 1$).	<ul style="list-style-type: none"> • Number partners for 2–5 using partner cards • Finding partners for numbers 2–5 • Using partner cards to model addition and subtraction • Connecting partner cards to combinations and story problems
4. For any number from 1 to 9, find the number that makes 10 when added to the given number, e.g., by using objects or drawings, and record the answer with a drawing or equation.	<ul style="list-style-type: none"> • Great Race game – object is to reach 10 • Making sticks of ten using Unifix cubes
5. Fluently add and subtract within 5.	<ul style="list-style-type: none"> • Number combinations using partner cards (2–5) • Number combination cards (add/subtract numbers 1–5)
Number and Operations in Base Ten	
<i>Work with numbers 11–19 to gain foundations for place value.</i>	
1. Compose and decompose numbers from 11 to 19 into ten ones and some further ones, e.g., by using objects or drawings, and record each composition or decomposition by a drawing or equation (e.g., $18 = 10 + 8$); understand that these numbers are composed of ten ones and one, two, three, four, five, six, seven, eight, or nine ones.	<ul style="list-style-type: none"> • Constructing numbers 1–19 with Unifix cubes (Magic Number) as a stick of 10 and so many more • Constructing numerals 11–19 using decade and unit cards
Measurement & Data	
<i>Describe and compare measurable attributes.</i>	
1. Describe measurable attributes of objects, such as length or weight. Describe several measurable attributes of a single object.	<ul style="list-style-type: none"> • Determine length of Unifix cube sticks by counting blocks
2. Directly compare two objects with a measurable attribute in common, to see which object has “more of”/“less of” the attribute, and describe the difference. For example, directly compare the heights of two children and describe one child as taller/shorter.	<ul style="list-style-type: none"> • Using Cardinality Chart, children determine more or less by the “height” of the dots
Geometry	
<i>Identify and describe shapes (squares, circles, triangles, rectangles, hexagons, cubes, cones, cylinders, and spheres).</i>	
1. Describe objects in the environment using names of shapes, and describe the relative positions of these objects using terms such as <i>above</i> , <i>below</i> , <i>beside</i> , <i>in front of</i> , <i>behind</i> , and <i>next to</i> .	<ul style="list-style-type: none"> • Line up objects on the number list and determine which is “in front” or “before” and which is “behind” or “after”

Common Core State Standards	Number Sense Intervention Activity
<i>Analyze, compare, create, and compose shapes.</i>	
3. Compose simple shapes to form larger shapes. <i>For example, “Can you join these two triangles with full sides touching to make a rectangle?”</i>	<ul style="list-style-type: none">• Join smaller sets of Unifix cubes to make a larger set

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Means and Standard Deviations by Group and Time (Raw Scores)

	Pretest		Post test		Delayed	
	M	SD	M	SD	M	SD
NSB Total						
Number Sense	17.02	6.80	31.74	8.13	33.00	7.96
Language	17.64	8.14	24.71	8.51	28.40	8.74
Control	19.64	8.61	25.27	8.27	29.30	8.38
NSB subarea						
Counting Skills						
Number Sense	5.93	1.99	7.57	2.03	7.88	1.73
Language	5.83	2.48	7.07	2.11	7.67	1.98
Control	5.86	2.34	7.50	1.92	7.98	1.76
Number Recognition						
Number Sense	2.36	1.64	6.19	2.12	6.48	1.48
Language	3.10	2.33	5.17	2.24	5.76	2.22
Control	3.55	2.55	5.41	2.23	6.27	2.06
Number Comparisons						
Number Sense	4.40	2.04	6.71	2.10	6.60	1.91
Language	4.24	2.00	5.48	1.99	6.07	1.81
Control	4.75	2.27	5.34	2.01	5.95	2.19
Nonverbal Calculation						
Number Sense	2.60	1.25	3.19	0.89	3.43	0.86
Language	2.57	1.27	3.19	1.04	3.43	0.77
Control	2.80	1.32	3.11	0.84	3.43	0.73
Story Problems						
Number Sense	0.95	0.99	3.98	1.42	3.62	1.58
Language	0.79	1.12	1.64	1.62	1.81	1.74
Control	1.05	1.14	1.64	1.64	1.98	1.62

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	Pretest		Post test		Delayed	
	M	SD	M	SD	M	SD
Number Combinations						
Number Sense	0.79	1.73	4.90	3.53	4.95	3.22
Language	1.12	1.81	2.55	2.90	3.71	3.12
Control	1.64	2.20	2.95	2.93	3.59	3.11
Vocabulary						
Number Sense	7.40	2.56	8.90	2.56	9.64	2.82
Language	7.67	2.72	10.12	3.07	10.86	2.87
Control	7.30	2.49	8.68	2.24	9.59	2.94
WJ Total						
Number Sense	9.64	3.30	16.67	5.19	17.07	4.93
Language	10.19	3.68	13.31	3.99	14.10	4.36
Control	10.50	4.01	13.50	3.41	15.34	4.18
WJ Applied Problems						
Number Sense	9.19	2.92	11.62	2.28	12.36	2.34
Language	9.29	2.71	11.00	2.63	11.07	2.65
Control	9.36	2.83	11.11	2.09	11.64	2.39
WJ Calculation						
Number Sense	0.45	0.97	5.05	3.81	4.71	3.40
Language	0.90	1.48	2.31	2.10	3.02	2.30
Control	1.14	1.86	2.39	2.10	3.70	2.41

Note. NSB= Number Sense Brief; WJ= Woodcock-Johnson; N = total number of items, M = Mean, SD = Standard deviation

Table 5

Results from the Multi-Linear Modeling (MLM) Analyses

Confirmatory Analyses				
	Number Sense vs. Controls		Language vs. Controls	
	<i>p</i> ¹	<i>d</i>	<i>p</i> ¹	<i>d</i>
NSB Total post	.01	1.80	<i>ns</i>	−0.21
NSB Total delayed	.01	0.63	<i>ns</i>	−0.18
WJ Total post	.01	0.91	<i>ns</i>	−0.15
WJ Total delayed	.01	0.83	<i>ns</i>	−0.17
Total Vocabulary post	.01	−1.44	<i>ns</i>	−0.19
Total Vocabulary delayed	.01	−1.07	<i>ns</i>	−0.21

Exploratory Analyses				
	Number Sense vs. Controls		Language vs. Controls	
	<i>p</i> ²	<i>d</i>	<i>p</i> ²	<i>d</i>
NSB Subarea				
Counting post	<i>ns</i>	0.11	<i>ns</i>	−0.22
Counting delayed	<i>ns</i>	0.03	<i>ns</i>	−0.22
Number recognition post	.011	0.61	<i>ns</i>	−0.26
Number recognition delayed	.029	0.49	<i>ns</i>	−0.22
Number comparisons post	.002	0.74	<i>ns</i>	−0.21
Number comparisons delayed	<i>ns</i>	.31	<i>ns</i>	−0.23
Nonverbal calculation post	<i>ns</i>	0.19	<i>ns</i>	−0.18
Nonverbal calculation delayed	<i>ns</i>	0.20	<i>ns</i>	−0.18
Story problems post	.001	2.64	<i>ns</i>	−0.23
Story problems delayed	.001	2.27	<i>ns</i>	−0.10
Number combinations post	.001	1.10	<i>ns</i>	−0.12
Number combinations delayed	.035	0.52	<i>ns</i>	−0.14
WJ Subtest				
WJ Applied Problems post	<i>ns</i>	.26	<i>ns</i>	−0.15
WJ Applied Problems delayed	.008	.64	<i>ns</i>	−0.15
WJ Calculation post	.001	1.48	<i>ns</i>	−0.16
WJ Calculation delayed	.008	1.00	<i>ns</i>	−0.20

Note: *d* is Cohen's (1988) effect size measure. Covariates in the model include pretest scores from both the Number Sense Brief and the WJMath Total.

¹ The overall error rate for confirmatory comparisons were adjusted using the Holm (1979) correction as recommended by the Institute of Education Sciences (IES) (cf. Schochet, 2008).

² As also per IES recommendations, the overall error rate was not adjusted for the exploratory comparisons.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 6

Percentile Rank Means and Standard Deviations on the WJ by Time and Group

	Pretest		Post test		Delayed	
	M	SD	M	SD	M	SD
Applied Problems						
Number Sense	33	25	40	21	41	25
Language	31	27	35	28	33	26
Control	36	22	36	19	36	22
Total	33	24	37	23	37	24
Calculation						
Number Sense	8	19	50	37	45	34
Language	18	26	27	33	31	32
Control	18	29	29	31	37	31
Total	15	25	35	35	37	33

Note. WJ = Woodcock-Johnson, M = Mean, SD = Standard deviation