Math Shelf: A Randomized Trial of a Prekindergarten Tablet Number Sense Curriculum

John Schacter, Jeff Shih, Charles M. Allen, Lina DeVaul, Amy B. Adkins, Taro Ito & Booil Jo

To cite this article: John Schacter, Jeff Shih, Charles M. Allen, Lina DeVaul, Amy B. Adkins, Taro Ito & Booil Jo (2015): Math Shelf: A Randomized Trial of a Prekindergarten Tablet Number Sense Curriculum, Early Education and Development, DOI: 10.1080/10409289.2015.1057462

To link to this article: http://dx.doi.org/10.1080/10409289.2015.1057462

Published online: 04 Sep 2015.

Article views: 17

View Crossmark data
Math Shelf: A Randomized Trial of a Prekindergarten Tablet Number Sense Curriculum

John Schacter\textsuperscript{a}, Jeff Shih\textsuperscript{b}, Charles M. Allen\textsuperscript{b}, Lina DeVaul\textsuperscript{b}, Amy B. Adkins\textsuperscript{b}, Taro Ito\textsuperscript{b} and Booil Jo\textsuperscript{c}

\textsuperscript{a}The Teaching Doctors, Stanford, California; \textsuperscript{b}College of Education, Center for Mathematics, Science and Engineering Education, University of Nevada, Las Vegas; \textsuperscript{c}Department of Psychiatry and Behavioral Sciences, Center for Interdisciplinary Brain Sciences Research, Stanford University

ABSTRACT

Research Findings: Effective preschool mathematics instruction is especially important for low-income children. Previous research demonstrates that low-income children enter kindergarten behind their middle-income peers. They receive less mathematics support at home and from public preschools. The aim of this study was to test Math Shelf, a tablet intervention designed to improve at-risk preschoolers' mathematics performance. A total of 100 children participated in a randomized controlled trial in a large urban Head Start center. Intervention students played Math Shelf on tablet computers for 6 weeks, whereas comparison students played the most downloaded and best reviewed preschool math apps on tablets for an equal amount of time. During game play, graduate student researchers supervised intervention and comparison students in separate rooms. Intervention and comparison groups did not differ on pretest assessments. Math Shelf students performed statistically significantly better (Cohen's $d = 0.57$) than comparison students at posttest. Practice or Policy: Math Shelf results suggest that teachers can enhance low-income preschoolers' mathematics knowledge in a relatively short amount of time by incorporating developmentally appropriate tablet interventions.


Socioeconomics-related mathematics differences

Socioeconomics-related math achievement differences are evident at home and in school (Alexander et al., 2001; Borman & Boulay, 2004; Klein & Starkey, 2004; Meyer, Princiotta, & Lanahan, 2004; Starkey & Klein, 2008). At home, middle- and high-income parents engage their children in more frequent and varied mathematical activities compared to low-income parents (Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010; Saxe, Guberman, & Gearhart, 1987; Starkey et al., 1999). Middle- and high-income parents use more mathematics language, read to their children more often, and watch less television than low-income families.

CONTACT: John Schacter, PhD \textsuperscript{a}schacter@teachingdoctors.com \textsuperscript{22 Pearce Mitchell Place, Stanford, CA 94305.}

Copyright © 2015 Taylor & Francis Group, LLC
do (Hofferth & Sandberg, 2001; Neuman & Celano, 2001; Starkey et al., 1999). During the summer, middle- and high-income children frequent libraries, bookstores, museums, zoos, aquariums, and historic sites more often than low-income children do (Meyer et al., 2004). Finally, middle- and high-income preschoolers’ parents have higher mathematics expectations compared to low-income preschoolers’ parents (Starkey & Klein, 2008).

Socioeconomics-related math achievement differences also occur in preschools (Early et al., 2005; Klein & Starkey, 2004). Fewer than 1 in 4 children who attend public preschools receive high-quality math instruction (Karoly, Ghosh-Dastidar, Zellman, Perlman, & Fernyhough, 2008). Public preschools spend less instructional time on math compared to private preschools, and public preschool students participate in fewer math activities than children in private preschools (Early et al., 2005; Rudd, Lambert, Satterwhite, & Zaier, 2008; Starkey & Klein, 2008). Qualitative studies have documented that teachers in low-income preschools use little mathematics language and report not knowing how to develop children’s number abilities (Bryant, Peisner-Feinberg, & Clifford, 1993; Farran, Silveri, & Culp, 1991; Rudd et al., 2008). Moreover, when public preschool teachers teach math, learning is hampered by ill-planned semiacademic activities (Stipek, Schoenfeld, & Gomby, 2012) in which the teaching of mathematics is secondary to other learning goals (National Research Council, 2009).

This lack of mathematics support in public preschools may explain low-income children’s poor mathematics performance in Head Start (Puma, Bell, Cook, & Heid, 2010). In a nationally representative sample of nearly 5,000 eligible preschoolers who were randomly assigned to Head Start or a control group, there was no statistically significant difference on math assessments at the end of Head Start, kindergarten, or first grade. Clearly, effective mathematics interventions for low-income preschoolers are necessary.

**Number knowledge development**

Developmental theory suggests that infants are born with the capacity to represent number in a nonverbal manner (Feigenson & Carey, 2003; Mix, Huttenlocher, & Levine, 2002). They can identify small numbers (i.e., less than 3), approximate larger number sets (Berch, 2005; Mix et al., 2002), and recognize transformations of small sets (Wynn, 1992). As infants become toddlers they acquire language and the ability to count. Counting extends children’s number understanding (Baroody, Lai, & Mix, 2006; Ginsburg, 1989).

Around the age of 4, children typically begin to merge their schemas for making global quantity comparisons with counting (Griffin, 2002; Jordan & Levine, 2009; Montessori, 1914). With teacher and parental guidance, children start to connect numerals with quantities, distinguish between successive numbers, and understand that numerals have magnitudes (Griffin, Case, & Siegler, 1994; Le Corre & Carey, 2007; Sarnecka & Carey, 2008; Schaeffer, Eggleston, & Scott, 1974). As these symbolic number knowledge skills develop, children begin to construct a mental number list (Frye et al., 2013; Siegler & Booth, 2004). Although most children reach these developmental milestones, a substantial number of students—typically those living in low-income communities—start kindergarten without them (Griffin, 2002; Jordan, 2007; Siegler, 2009).

**Teaching number knowledge skills**

To improve low-income preschoolers’ number knowledge skills, researchers have developed a variety of successful interventions (Baroody, Eiland, & Thompson, 2009; Chard et al., 2008; Griffin et al., 1994; Jordan, Glutting, Dyson, Hassinger-Das, & Irwin, 2012; Ramani & Siegler, 2008). Ramani and Siegler (2008) improved low-income 4-year-olds’ number knowledge skills by playing linear numerical board games. In just four 15-min sessions, children who played
board games with consecutively numbered, linearly arranged, equal-size spaces performed statistically significantly better on number line estimation tasks than children who played an identical board game but with colors instead of numbers on the game board. According to the authors, moving a token in such a game teaches children number names, number magnitude, and number sequence because the distance from the start of the board increases with each physical move.

Baroody et al. (2009) improved low-income preschoolers’ number knowledge by using manipulatives to teach verbal counting, object counting, and numeral–quantity relationships. Nancy Jordan and colleagues have developed several successful number knowledge interventions for low-income children (Dyson, Jordan, & Glutting, 2013; Jordan et al., 2012). They have implemented numerical board games similar to Ramani and Siegler (2008), implemented number before and after activities on a number line like Chard et al. (2008), and created original games like the Magic Number 10 to introduce place value, the hundreds chart to teach numeral sequencing, and counting strategies to solve story problems.

Finally, Sharon Griffin’s (2004a,b) effective number knowledge curriculum, Number Worlds, exposes preschoolers to quantities, counting, and formal symbols and provides multiple opportunities for constructing relationships among these three ways of understanding number. Like other successful preschool number sense interventions, Griffin uses games that teach subitizing, counting aloud, numeral identification, numeral sequencing, number magnitude, and matching numerals to quantities. As children play with Griffin’s five different forms of number representations (i.e., groups of objects, dot patterns, positions on a horizontal line, positions on a vertical line, and position on a dial), they develop foundational skills to connect these number representations to their number experiences in the world.

The wide variety of number knowledge skills taught in these effective interventions is often referred to as number sense in the research literature. Number sense describes a child’s flexibility and fluidity with quantities and numbers (Griffin, 2002). Specifically, number sense includes the recognition that numerals represent quantities and have magnitudes and that sets of numbers can be transformed and manipulated (Gelman & Gallistel, 1978; Griffin, 2004b; Malofeeva, Day, Saco, Young, & Ciancio, 2004).

Tablets to advance number learning

Thus far, effective number sense interventions have used people (teachers, researchers, graduate students) to teach number knowledge skills to preschoolers. Tablet computers can also provide targeted number sense learning opportunities. Young children find tablets highly motivating (Flewitt, Messer, & Kucirkova, 2014) and learn how to use them almost immediately (Boddu, 2013). Tablet touch screens allow children with limited fine motor skills to operate these devices with their fingers, thereby eliminating the more complex hand–eye coordination required to use a keyboard and mouse (L. Z. Cooper, 2005). Apps that introduce learning games, puzzles, and videos have become a part of many middle- and high-income children’s daily lives (Flewitt et al., 2014; Schneider et al., 2012; Siegle, 2013).

Although there is limited evidence of the efficacy of using tablets to increase preschoolers’ mathematics achievement, recent meta-analytic reviews of computer game-based learning (Ke, 2009; Wu, Choiu, Kao, Hu, & Huang, 2012) have shown that games designed based on learning and developmental theories can enhance student achievement. Moreover, teachers who think about ways to appropriately incorporate tablets and computer games into their curriculum are more likely to increase students’ learning than educators who reject these technologies outright (Chai, Ng, Li, Hong, & Koh, 2013; Clarke, Svanaes, & Zimmermann, 2013).
Math Shelf, a tablet number sense intervention

Like other effective number sense interventions, Math Shelf relies on developmental research to guide the math skills and sequence of skills that are introduced (Clements & Sarama, 2009; Jordan & Levine, 2009; Siegler, 2009). Math Shelf includes a variety of games, puzzles, and manipulatives to maintain young children’s interest and foster their engagement. What differentiates Math Shelf from other research-based number sense interventions is that all instruction and practice occurs on tablet computers.

Following Jordan and Levine’s (2009) recommendations, Math Shelf games first teach the quantities 1 to 5, focusing on subitizing, ordering quantities, one-to-one counting, and matching different quantity representations (see Figure 1). Four games teach these skills (12 activities per game), each with a different virtual manipulative (i.e., colored beads, dot cards, counters, and number rods). The first of these games focuses on counting, subitizing, and matching different quantities from 1 to 3.

Figure 2 demonstrates how the concepts of connecting number names to symbols, ordering numerals, matching numerals to quantities, and counting to apply the cardinal principle are introduced.

When children demonstrate mastery of the numerals and quantities 1 to 5 (i.e., completing 48 activities at an 80% correct level), the same skills with new games and activities are practiced with the numerals and quantities 1 to 9. Finally, numeral and collection comparisons are introduced to teach number magnitude (see Figure 3).

![Figure 1. Colored Beads 1 to 5. Panel 1 directs the child to sequence the beads least to greatest. Panel 2 poses a series of questions that ask the child to touch the quantity that corresponds to the named number without counting (subitizing). Panel 3 requires the child to match the bead quantity to the finger quantity representation by sorting.](image1)

![Figure 2. Dot Cards 1 to 5. Panel 1 teaches number identification using a jumping game. Panel 2 requires students to order numbers 1 to 5, then to match numbers to quantities. Panel 3 has children count the animals (in this example ducks) and touch the last number counted (cardinality).](image2)
Math Shelf incorporates three effective feedback strategies identified in Hattie and Timperley’s (2007) synthesis of 12 meta-analytic studies. First, Math Shelf’s feedback relates directly to the learner’s goal. Second, Math Shelf’s feedback provides cues and reinforcement. Third, feedback is computer assisted to show students how to correctly complete the task.

Each Math Shelf activity presents a clear performance goal. For example, several Math Shelf tasks require sorting different virtual manipulatives (number rods, colored beads, dot cards) from least to greatest (see Figure 4). Before the child begins, the correctly sorted manipulatives are displayed, demonstrating the performance goal (see Panel 1). Next, the manipulatives are mixed up, and the student must resequence them to achieve the goal. To assist the child, Math Shelf provides two cues: (a) a global outline of the sequence structure and (b) an animated finger that shows the student to drag the first manipulative into its correct position (see Panel 2). Children can choose to follow these cues or to drag and drop the manipulatives into incorrect positions. When all virtual manipulatives are placed in the global outline, the incorrect manipulatives are returned along with an audio sound that indicates that some were sequenced incorrectly. If the student is unsuccessful three times, a third cue (i.e., individual outlines of each manipulative) shows the student where to drag and drop each number rod (see Panel 3). After all manipulatives are correctly placed, audio feedback (a soft chime) reinforces the learner.

Along with providing goal-directed feedback, cues, and reinforcements, Math Shelf also records each child’s performance in a database. Tasks on which children are unsuccessful three times are programmed to reappear in that child’s future game play sequence. This ensures that each student receives additional practice until he or she is able to complete the activity without scaffolds.

**Math Shelf’s feedback strategies**

Figure 3. Number Rods 1 to 9. Panel 1 asks children to sort the number rods least to greatest, then drag the flying numbers to the correct rods. In Panel 2, children must match the rod quantity to the number in order to color the picture. Panel 3 asks children to touch either the smallest or largest number rod. As this activity progresses, children must determine magnitude with only numbers.

Figure 4. Math Shelf feedback. Panel 1 shows the performance goal (i.e., to sequence the rods from least to greatest). In Panel 2, a pointed finger provides a cue to drag the 1 rod into the first position. If a child sorts the rods incorrectly three times, outlines along with a pointed finger show where to drag and drop each rod (Panel 3).
The present study

Large mathematics achievement differences between low-income and middle-income children are present when these children enter preschool and kindergarten (Ginsburg & Russell, 1981; Hughes, 1986; Jordan et al., 1992; Jordan & Levine, 2009; Klein et al., 2008). Math Shelf’s purpose is to develop low-income preschoolers’ number sense to prepare them for kindergarten and beyond. Kindergarten Common Core State Standards for mathematics expect children to count to 100 by tens and ones, understand the relationship between numerals and quantities, count on from a given number, compose and decompose numbers from 11 to 19 into tens and ones, and add and subtract numbers within 10. Children who enter kindergarten with number sense knowledge will be more likely to achieve these standards and succeed in later grades because early number sense knowledge is highly predictive of mathematics achievement through at least third grade (Jordan, Glutting, & Ramineni, 2010; Jordan, Kaplan, & Olah Nabors, 2006; Jordan, Kaplan, Ramineni, & Locuniak, 2009). Moreover, preschool mathematics ability is a positive and statistically significant predictor of mathematics achievement through age 15, even after differences in other academic skills, attention, and personal and family background characteristics are adjusted (Watts, Duncan, Siegler, & Davis-Kean, 2014).

Research questions

1. Will low-income children who play Math Shelf perform better on number sense assessments than similar students who play with the five best selling and best reviewed prekindergarten (pre-K) mathematics apps sold on iTunes in 2014?
2. Will children assigned to Math Shelf, regardless of their initial number sense knowledge, exhibit similar learning growth?
3. Will girls and boys benefit equally from the Math Shelf intervention?

Method

Participants

A total of 100 preschool children from a large Head Start center in Nevada were randomly selected (from 164 total students across eight classrooms) to participate in the study. Of these, 50 students were randomly assigned to the intervention and 50 to the comparison group. At the pretest date, the sample consisted of 50 intervention and 50 control group students. At the posttest date, 45 intervention and 41 comparison children remained. The attrition rate was lower than the monthly attrition reported by the center of between 24% and 36%. Participants’ characteristics are stated in Table 1.

<table>
<thead>
<tr>
<th>Baseline Characteristic</th>
<th>Intervention (n = 50)</th>
<th>Comparison (n = 50)</th>
<th>Group Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest, M (SD)</td>
<td>21.7 (11.8)</td>
<td>21.4 (11.4)</td>
<td>t(99) = 0.15, p = .89</td>
</tr>
<tr>
<td>Age in months, M (SD)</td>
<td>56.3 (7.9)</td>
<td>55.4 (7.2)</td>
<td>t(99) = 0.64, p = .53</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>23 (46)</td>
<td>25 (50)</td>
<td>χ²(1) = 0.16, p = .69</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td>χ²(3) = 0.22, p = .97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic 23 (46)</td>
</tr>
<tr>
<td>Black 21 (42)</td>
</tr>
<tr>
<td>White 5 (10)</td>
</tr>
<tr>
<td>Multiracial 1 (2)</td>
</tr>
</tbody>
</table>
Procedures

Children assigned to the intervention played the Math Shelf games on iPads. Students assigned to the comparison group had the choice to play the five most downloaded and best reviewed 2014 pre-K math apps on iPads. Game play for the treatment and comparison conditions occurred in separate classrooms supervised by graduate students. This arrangement prevented interaction across randomized experimental conditions and limited the threat to internal validity due to contamination of this randomized experiment (Rubin, 1978, 1980, 1990).

Children in the intervention and comparison conditions played 3 days a week, for 10 min a session, for 6 weeks from April 28, 2014, to June 6, 2014. Graduate students supervised three children playing at a time, each on their own iPads for 10-min intervals. Each iPad was connected to headphones to reduce noise disruptions.

Intervention software. As described previously, Math Shelf is a tablet number sense intervention. Developmental research informed the number knowledge skills and sequence Math Shelf introduced (Clements & Sarama, 2009; Jordan & Levine, 2009; Siegler, 2009). In addition, Math Shelf utilized research from successful early number sense interventions to guide content and sequencing choices (Baroody et al., 2009; Chard et al., 2008; Clements & Sarama, 2009; Griffin, 2004b; Jordan et al., 2012; Ramani & Siegler, 2008).

Math Shelf transforms three of Maria Montessori’s physical manipulatives (colored beads, dot cards/counters, and number rods) into virtual manipulatives. These virtual manipulatives are used in a variety of games, puzzles, and activities to teach subitizing, counting (one-to-one and cardinality), sequencing quantities, numeral identification, matching numerals and quantities, sequencing numerals, and comparing quantity and numeral magnitudes. Four games (12 activity screens per game) teach these skills for the quantities and numerals 1 to 5 (the first of these four games teaches only the quantities 1 to 3). Five games (10 activity screens per game) teach these skills for the quantities and numerals 1 to 9. Thus, children played a total of 98 unique activities.

Prior to the randomized trial being conducted, Math Shelf was developed and tested for 12 months in a Head Start center in northern California. During development, games were revised repeatedly based on observing children play and conducting informal interviews with students and teachers. Along with implementing developmental and early intervention research, Math Shelf’s numerous design iterations helped to ensure that the content, sequence of skills, and difficulty level of activities were developmentally appropriate before the experiment began.

Lastly, Math Shelf administered a 12-item placement test to determine whether each child should start at the beginning (i.e., the first game, which taught number concepts 1 to 3) or at the midpoint (i.e., the game that began teaching number concepts to 9) of the intervention. The placement test asked children to order numerals from 1 to 6, then match quantities to numerals 1 to 6 (see Figure 5). Children who answered eight of 12 items correctly started Math Shelf learning number concepts to 9. A total of 18% of treatment children (i.e., nine students) started the intervention playing the 1-to-9 games. The remaining 41 students began the intervention playing the first Math Shelf game that taught number concepts 1 to 3. Of the nine children who began the intervention playing the 1-to-9 games, two completed all games during the fifth week of the study. These two children were instructed to select the games and activities they enjoyed most and replay them during Week 6, the final week of the intervention.

Control software. Teachers who incorporate tablets into their instruction download apps from iTunes or Google Play. In order to vet apps for quality, teachers may rely on reviews of apps published by reputable sources and the top-grossing app lists in the education category. Both of these

1Math Shelf© was created, designed, and developed by John Schacter, PhD.
sources were used to select five pre-K math apps for the comparison group children to play on iPads. We first selected the two top-grossing pre-K math apps in 2014: (a) Team Umizoomi and (b) Numbers With Nemo. Next we chose three of the most widely and best reviewed pre-K math apps by Children’s Tech Review, Moms With Apps, the Parent Choice Awards, Common Sense Media, and USA Today’s Top 10 Apps for Kids. The three best reviewed apps were Monkey Math, Elmo Loves Math, and Park Math HD. These apps taught various pre-K and kindergarten math skills and content.

**Test administration procedures.** Graduate student researchers tested all children individually on an iPad number sense assessment that provided audio and visual instructions. All children were pretested during the week of April 21, 2014, and posttested during the week of June 9, 2014. Test administration scripts were strictly followed.

**Measures**

A 62-item iPad number sense assessment was developed for the study. The dependent measure included constructs that assessed goals recommended by numerous early education researchers who study number sense. Moreover, the Math Shelf intervention software was directly aligned to these recommended number sense concepts by the dependent measure. The iPad number sense assessment included numeral identification and quantity discrimination measures studied by both Lee, Lembke, Moore, Ginsburg, and Pappas (2007) and Jordan and associates (2012). The assessment also incorporated number–object correspondence and comparing quantity tasks created by Malofeeva and colleagues (2004) as part of their number sense assessment for Head Start preschoolers. Finally, the iPad number sense test used numeral sequencing items similar to those in Seethaler and Fuchs’s (2010) early education number sense battery.

The untimed iPad number sense assessment took children approximately 6 min to complete. Test–retest reliability was collected on a sample of 20 students (average age = 4 years, 5 months) who took the test 5 days apart in a Head Start center in northern California. The test–retest reliability intraclass correlation was .97. Cronbach’s alpha inter-item reliability was .94. The iPad number sense assessment tested the following concepts.

**Quantity discrimination (six items).** Students were presented with four random numerals (ranging from 1 to 10) and asked to touch the largest numeral. All children received the same four random numerals in the same order. The problems were as follows: (a) 1, 3, 7, 4; (b) 8, 2, 9, 5; (c) 2, 10, 3, 6; (d) 4, 2, 5, 1; (e) 8, 4, 3, 6; and (f) 7, 3, 6, 10. Clarke and Shinn (2004) reported

**Numerical identification (eight items).** Four numerals from 1 to 10 were displayed, and students were instructed to touch one of the numerals. The four numerals were presented in random order, with correct answers appearing in different positions each time. Lee et al. (2007) reported a Cronbach’s alpha coefficient of .88 and concurrent criterion validity of .59 with the Test of Early Mathematics Ability–3 (Ginsburg & Baroody, 2003) for numeral identification tasks. Clarke and Shinn (2004) documented predictive validity of .68 with the Woodcock–Johnson spring first-grade Applied Problems subtest (Woodcock & Johnson, 1989).

**Numerical sequencing (16 items).** In the first task, the numerals 1 to 6 were displayed in random order, and children were asked to sequence them from smallest to largest. A second task that occurred later in the assessment had children sequence the numerals 1 to 10. Seethaler and Fuchs (2010) used a similar numeral sequencing task as part of their number sense battery. Reliability was .70, with predictive validity of .67 for the total math score in first grade on the Iowa Test of Basic Skills.

**Cardinal principle (eight items).** This series of items assessed the child’s knowledge of counting and the cardinal principle. That is, after counting a group of animals on the iPad, the child was asked to identify the numeral representing how many total animals were in the set. These items were similar to nine items Malofeeva et al. (2004) developed called Number-Object Correspondence Task 3. Cronbach’s alpha reliability was .86.

**Comparing quantities (eight items).** In this task children compared four different quantity collections represented as dot cards from 1 to 9, then selected either the smallest or largest. Griffin and colleagues (1994) developed a similar task in their number sense assessment. Sasanguie, Van den Bussche, and Reynvoet (2012) reported that comparing quantities predicted a large and statistically significant amount of variance in children’s future math achievement.

**Matching numerals to quantities (16 items).** Kindergarten Mathematics Common Core State Standard B.4 states that children should “understand the relationship between numbers and quantities” (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p. 10). To assess this skill, the iPad showed students various dot card quantity representations and asked them to match each quantity to the correct numeral.

**Results**

To examine the baseline balance between the two randomized groups on the number sense assessment, we used two-sample *t* tests (for continuous variables) and chi-square tests (for categorical variables). Table 1 shows that there were no statistically significant differences between the intervention and comparison groups in terms of number sense pretest, age, gender, or race. The finding of no statistically significant differences between the treatment and control groups on baseline characteristics supports the conclusion that randomization was successful.

Standard longitudinal linear mixed effects modeling (e.g., Singer & Willett, 2003) was used to test the change from the pre- to postintervention. Specifically, we used a random intercept model assuming a linear trend over time. We chose mixed effects modeling instead of conventional regression or analysis of covariance to better handle missing data and to stay in line with the intention-to-treat principle (Little & Rubin, 2002). We included all individuals randomized in the analyses for whom data were available from at least one of the two assessments (pre and/or post).
Data points that were not available were treated as missing at random conditional on observed information using a maximum likelihood estimation (Little & Rubin, 2002). For the Maximum Likelihood Expectation Maximization estimation of our models, we used Mplus Version 7.3. All 100 individuals randomized either to the intervention \((n = 50)\) or to the comparison condition \((n = 50)\) had a pretest score and were included in the mixed effects modeling. At the time of the posttest, five dropouts were assigned to the intervention and nine dropouts were assigned to the comparison group.

Table 2 summarizes the results of our mixed effects modeling. Students in both conditions improved. There was a statistically significant and sizable effect for the intervention on number sense compared to the control group \((\text{Cohen’s } d = 0.57, p < .001)\). Given that we only had eight classrooms, we did not account analytically for possible classroom effects, which may have resulted in a somewhat inflated Type I error rate. However, our primary group comparison reported in Table 2 showed a robust result \((p < .001)\), and therefore the conclusion is unlikely to change unless the classrooms are unusually heterogeneous.

In exploratory analyses, we examined potential moderators of the intervention’s effect on number sense. The analytical criteria used for detecting moderators conformed to the MacArthur approach (Kraemer, Wilson, Fairburn, & Agras, 2002), which was embedded in the mixed effects modeling framework. As described in Table 1, all four variables (pretest, age, gender, and race) showed no difference across the intervention and control groups, satisfying the eligibility criteria for moderators.

In our mixed effects modeling framework, the key parameter of interest was the effect of the interaction between the intervention status and a potential moderator on the improvement in number sense. If this parameter estimate was statistically significant, the baseline variable satisfied the analytical criteria for moderators and therefore was identified as a moderator. Among the four baseline variables examined as potential moderators, pretest and gender were found to be intervention effect moderators (see Table 3). The effect of the intervention on

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some prior number knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>32.1</td>
<td>51.0</td>
</tr>
<tr>
<td>Control</td>
<td>32.1</td>
<td>35.0</td>
</tr>
<tr>
<td>Minimum prior number knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>12.9</td>
<td>22.3</td>
</tr>
<tr>
<td>Control</td>
<td>12.9</td>
<td>17.5</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>21.1</td>
<td>35.0</td>
</tr>
<tr>
<td>Control</td>
<td>21.1</td>
<td>22.1</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>22</td>
<td>34.2</td>
</tr>
<tr>
<td>Control</td>
<td>22</td>
<td>29.2</td>
</tr>
</tbody>
</table>
number sense was greater ($p = .004$) for those who started with higher pretest scores (upper 50%, pretest $\geq 20$). The effect of the intervention was also greater for female students ($p = .034$).

**Discussion**

The results from this randomized trial show that low-income children who played Math Shelf performed statistically significantly better on a number sense assessment compared to preschoolers who played the best reviewed and most downloaded pre-K math apps of 2014. Applying Hill, Bloom, Black, and Lipsey’s (2007) effect size interpretation for normative change, we find that the group difference effect size of 0.57 can be translated as Math Shelf students being 6 months ahead of the comparison group in terms of their number sense development.

Our results support previous early intervention number sense research by demonstrating that children who are economically disadvantaged benefit from early number sense interventions (Baroody et al., 2009; Chard et al., 2008; Griffin, 2004b; Jordan et al., 2012; Ramani & Siegler, 2008). Our results add to the literature by showing that tablet games designed based on developmental theory can enhance low-income preschoolers’ number sense knowledge.

Although all Math Shelf children made greater number sense progress than control students, some intervention children learned more than others. Exploratory analyses suggested that preschoolers with little to no number sense knowledge learned less than children who began the intervention with some number sense understanding. Furthermore, girls appeared to learn more from the intervention than boys.

For preschoolers with little to no number sense knowledge, Math Shelf games, puzzles, and activities may not have provided enough practice on their developmental level. These children began Math Shelf by playing 12 activities that taught 1-to-3 subitizing, counting and cardinality, sequencing, and matching different representational collections. Next they participated in activities that taught these skills from 1 to 6, in addition to new skills such as numeral identification, numeral and quantity sequencing, and number magnitude. Most likely, Math Shelf advanced students with little to no number sense knowledge too quickly. Research suggests that more practice subitizing 1 to 3 would likely increase competence and form a stronger foundation for meaningful learning of early number concepts (Baroody et al., 2006; Benoit, Lehalle, & Jouen, 2004; Butterworth, 2005; Hannula, Lepola, & Lehtinen, 2010).

Based on this exploratory finding, the Math Shelf development team has created 40 new games and activities focused solely on 1-to-3 number concepts. This additional practice subitizing, counting, sorting, sequencing, and matching quantity representations from 1 to 3 will likely better prepare children with little to no number sense to be successful on more advanced number sense tasks.

The exploratory finding that girls made greater mathematics progress than boys is surprising. Previous meta-analytic reviews of computer-assisted instruction in elementary schools have not demonstrated gender effects (Blok, Oostdam, Otter, & Overmaat, 2002; Fletcher-Finn & Gravatt, 1995; Kulik, Kulik, & Bangert-Drowns, 1985; Weng, Maeda, & Bouck, 2014). Examination of the means shows that growth was similar in the intervention for both groups, but boys benefited more from the comparison software. Perhaps the Math Shelf software is more gender neutral than the control group software. Or the Math Shelf virtual manipulatives, puzzles, games, and characters appealed equally to boys and girls. Alternatively, because the sample size in the exploratory analysis was relatively small (i.e., 20 boys and 24 girls), this finding may be a spurious result. Future research with larger numbers of participants and more detailed ethnographic notes may help determine whether this result persists and why.

**Study limitations**

The first limitation of this study is that the intervention and control conditions were not integrated into each classroom’s mathematics curriculum and teacher training. To allow for causal
interpretations of the Math Shelf intervention, our research design had children play the treatment and comparison tablet games in different rooms under the supervision of graduate students. Although this arrangement limited the threat to internal validity due to contamination (Rubin, 1978, 1980, 1990), it reduced the ecological validity of the study and may have resulted in stronger findings than would be expected under less controlled circumstances. We are currently planning to study Math Shelf in classroom contexts. Teachers will be trained in how to organize their rooms and schedule and monitor children playing Math Shelf 2 days a week. Each room will use three iPads, a timer, and a large poster with students’ names and game play schedules to regularly implement the intervention. This new study design will better assess the ecological validity of Math Shelf.

The second limitation of this research concerns the collection of fidelity of implementation data for both intervention and control groups. Although treatment and control students played the games on tablets 3 days per week, 10 min each day, researchers did not keep detailed records of which students in each condition were absent. In future studies better fidelity information can serve as an additional moderator for exploratory analyses. The third limitation of this study is the fact that we did not analytically account for possible classroom effects (i.e., an eight classroom sample size is too small for accurate analyses), which may have resulted in an inflated Type I error rate.

Lastly, this study did not implement a delayed posttest to measure whether the number sense knowledge gained from playing Math Shelf persisted over time. Although other researchers have shown the lasting benefits of early number sense knowledge, a delayed posttest would provide direct evidence for lasting intervention results.

**Conclusion**

Increasing low-income preschoolers’ mathematics achievement has been a very difficult endeavor. Of 15 research-based pre-K mathematics curricula evaluated in the Preschool Curriculum Evaluation Research Consortium’s (2008) *Effects of Preschool Curriculum Programs on School Readiness*, only one curriculum improved students’ math achievement, and that program combined computer-based math instruction with teacher-directed hands-on activities and parent training (Klein et al., 2008).

The importance of early number sense knowledge to future mathematics achievement is clear (Jordan et al., 2006, 2009, 2010). This randomized experiment provides initial evidence that Math Shelf increases low-income children’s number sense knowledge. By placing greater emphasis on evidence-based number sense programs, public preschools can truly provide a head start for low-income children’s mathematics achievement.

**Acknowledgments**

We thank Angel Barrios, Rocky Sodetani, Rory Sipp, and Pedro Castellanos for their hard work coordinating this research. We especially thank Dr. Sandra Twardosz and the five reviewers of this article for their guidance and insights.

**References**


