Research Reports

Domain-Specific and Domain-General Training to Improve Kindergarten Children’s Mathematics

Geetha B. Ramani*, Susanne M. Jaeggib, Emily N. Dauberta, Martin Buschkuehlc

[a] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, USA. [b] School of Education, University of California, Irvine, CA, USA. [c] MIND Research Institute, Irvine, CA, USA.

Abstract

Ensuring that kindergarten children have a solid foundation in early numerical knowledge is of critical importance for later mathematical achievement. In this study, we targeted improving the numerical knowledge of kindergarteners (n = 81) from primarily low-income backgrounds using two approaches: one targeting their conceptual knowledge, specifically, their understanding of numerical magnitudes; and the other targeting their underlying cognitive system, specifically, their working memory. Both interventions involved playing game-like activities on tablet computers over the course of several sessions. As predicted, both interventions improved children’s numerical magnitude knowledge as compared to a no-contact control group, suggesting that both domain-specific and domain-general interventions facilitate mathematical learning. Individual differences in effort during the working memory game, but not the number knowledge training game predicted children’s improvements in number line estimation. The results demonstrate the potential of using a rapidly growing technology in early childhood classrooms to promote young children’s numerical knowledge.

Keywords: mathematics, working memory, achievement gap, intervention

Providing children with a strong foundation in mathematics is of critical importance since early mathematics performance is highly predictive for later mathematical achievement (Duncan et al., 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009). By the time children reach 12th grade, 75% of students score below “Proficient” on standardized tests (National Center for Education Statistics, 2015). Beyond high school, 58% of American adults cannot compute a 10% tip and 71% cannot calculate miles per gallon on a trip (Phillips, 2007). Having poor mathematical skills has consequences for daily functioning and career advancement as strong mathematics knowledge is increasingly important in our technological society (Kilpatrick, Swafford, & Findell, 2001). A serious concern is that deficits in numerical knowledge are found before children even enter kindergarten. Furthermore, there are large disparities between children from lower- and higher-SES backgrounds (Jordan, Kaplan, Oláh, & Locuniak, 2006; Starkey, Klein, & Wakeley, 2004), and these early gaps widen across school years (McClelland, Acock, & Morrison, 2006). For these reasons, fostering children's
mathematical understanding at school entry, especially in children from low SES backgrounds, is crucial for ensuring students’ success in school and beyond.

The current study aimed to improve kindergarten children’s numerical knowledge using tablet computer games. We tested the specificity of training activities on either domain-specific or domain-general skills as well as individual factors that may influence children’s learning from these activities. We focused primarily on children from low-income backgrounds since their numerical knowledge tends to trail behind children from middle-income backgrounds, and because they are most at risk for developing mathematical learning difficulties (Jordan & Levine, 2009).

Conceptual Framework of Mathematical Achievement

According to Geary and Hoard’s (2005) framework on mathematics, both numerical knowledge and processing abilities contribute to mathematical achievement. Specifically, children’s conceptual and procedural knowledge contribute to their competencies in an area of mathematics, such as arithmetic (Rittle-Johnson & Siegler, 1998; Rittle-Johnson, Siegler, & Alibali, 2001). Children’s conceptual knowledge of the magnitudes of numbers is especially important because it involves the ability to estimate the magnitude of numerals and sets of objects (Siegler, 2016; Siegler & Booth, 2005). This conceptual knowledge can impact children’s use and success of their procedural knowledge in arithmetic. Empirical evidence demonstrates the importance of numerical magnitude knowledge. For example, children’s performance on measures of numerical magnitude knowledge correlates strongly with mathematical achievement test scores from grades ranging from kindergarten through sixth grade (Booth & Siegler, 2006; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Schneider, Grabner, & Paetsch, 2009). Further, children with more advanced numerical magnitude knowledge in the first grade show faster growth in mathematical skills over the elementary school years (Geary, 2011).

The model also specifies that deficits in conceptual and procedural knowledge could be due to deficits in underlying cognitive systems, specifically working memory (WM). WM is an essential system that underlies the performance of virtually all complex cognitive activities including mathematics learning (Diamond, 2013; Shah & Miyake, 1999). People differ in how much information they can hold in WM, and how well they can hold that information in the face of distraction (Engle, Kane, & Tuholski, 1999; Jonides et al., 2008). These individual differences can play a central role in scholastic achievement (Pickering, 2006).

Individual differences in underlying cognitive systems, specifically, WM, predict performance in mathematics tasks requiring both, conceptual and procedural knowledge (see Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Raghubar, Barnes, & Hecht, 2010 for recent a recent review and meta-analysis). For example, WM plays a substantial role in the development of numerical magnitude knowledge (Kolkman, Hoijtink, Kroesbergen, & Leseman, 2013). Specifically, number line estimation, which is a common way to assess children’s magnitude knowledge, involves showing children lines with a number at each end (e.g., 0 and 100) and a third number, printed above the line, within the given range. The task is to estimate the location of the third number on the line (e.g., “Mark where 36 would go on the line”). To complete this task, children must first identify and encode the Arabic numeral to estimate. They must then process the magnitude of the given number, interpret its magnitude in relation to the numbers at each end of the line, while simultaneously comparing the information in different representations (e.g., Arabic numerals and visual representation of the
number line). Given these processing demands, deficits in WM might be among the restricting factors for children’s ability to complete the necessary steps of this task and estimate the numerical magnitude.

Similarly, WM is critical for children’s procedural knowledge and execution of the action sequences used for solving arithmetic problems (Geary, Hoard, Byrd-Craven, & DeSoto, 2004). For example, poor WM can lead children to struggle with more advanced forms of arithmetic strategies, such as counting on from the larger number during addition, because they have difficulty keeping track of the numbers. As a result, children will use less sophisticated arithmetic strategies, which may ultimately restrict the acquisition of more complex strategies (Geary & Hoard, 2005).

**Numerical Knowledge, Working Memory, and Socioeconomic Status (SES)**

The mathematical performance of young children from lower-income backgrounds is on average lower than that of children from higher-income backgrounds. This achievement gap can widen over the course of schooling (Alexander & Entwisle, 1988). These early differences in numerical knowledge are primarily found on verbal and symbolic number skills, such as tasks with verbally stated numbers, story problems, and identifying written numbers (Demir, Prado, & Booth, 2015; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Starkey, Klein, & Wakeley, 2004). SES-related differences have also been found on tasks assessing children’s numerical magnitude knowledge (Siegler & Ramani, 2008). This suggests that SES-related differences are specific to symbolic number skills, which are critical for higher-level mathematics, and are also highly dependent on input, verbal ability, instruction, and experience (Jordan & Levine, 2009).

Children from low-income backgrounds are also vulnerable to develop deficits in domains that rely on the integrity of frontal brain regions, such as WM and related skills (Davis-Kean, 2005). Differences in WM between children from lower- and higher-income backgrounds have been observed in first grade children and adolescents in the U.S (Finn et al., 2016; Noble, McCandliss, & Farah, 2007), as well as in even younger children in other countries (Fernald, Weber, Galasso, & Ratsifandrihamanana, 2011; Lipina et al., 2013). These SES-related deficits in WM may be exacerbating the gap in mathematical achievement between children from higher-income backgrounds and children from lower-income backgrounds. In sum, deficits in both numerical knowledge and WM are observed in children from low SES-backgrounds, and it is critical to address any deficits at an early age in order to mitigate the associated long-term consequences.

**Domain-Specific Training of Numerical Magnitude Knowledge**

Research has shown that playing informal learning activities, specifically linear number board games, can promote children’s numerical magnitude knowledge. Linear number board games are physical representations of the mental number line, which is hypothesized to represent numbers horizontal from left to right in Western cultures (for reviews, see Ansari, 2008; Hubbard, Piazza, Pinel, & Dehaene, 2005; Siegler 2016). Consistent with this view, Laski and Siegler (2014) have proposed the cognitive alignment framework suggesting that learning materials are more beneficial when they more closely align with this desired linear mental representation.

Empirical work supports this approach and has shown that playing linear numerical board games, but not circular numerical board games, can improve young children’s numerical knowledge. In several studies, preschoolers were randomly assigned to either play a linear numerical board game with squares numbered...
from 1 to 10 or an identical game, except for the squares varied in color rather than number. After four 20 minute sessions, the children who played the numerical version of the game showed greater improvements in number line estimation, magnitude comparison, counting, numeral identification, and ability to learn novel arithmetic problems (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009). These improvements were stable for at least a two-month period (Ramani & Siegler, 2008). However, playing board games for the same amount of time, but with the numbers arranged circularly, did not produce the same improvements (Siegler & Ramani, 2009). Other studies using number board games have replicated and extended this previous research with preschool children in Scotland (Whyte & Bull, 2008).

Promising results have also been found using computer software to improve children’s numerical knowledge. For example, children from lower-SES backgrounds between the ages of 4-6 played an adaptive computer game, “The Number Race,” that focused on performing numerical comparisons using dots, numbers, or arithmetic problems for six 20-min sessions. Playing the mathematics game improved children’s numerical magnitude knowledge, whereas playing a reading game did not (Wilson, Dehaene, Dubois, & Fayol, 2009). In a follow-up study with kindergarten children, using the same number comparison game, as well as a game focused on mapping sets of objects to numbers, children’s numerical comparison improved after playing either of the two games. However, improvements were not found in the other areas of numerical knowledge tested, such as counting and arithmetic (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009). A more recent study found similar results with middle-income 4- to 6-year-olds, who played “The Number Race,” which improved children’s mental arithmetic, numerical magnitude knowledge, and their numeric semantic knowledge, such as mapping digits to a corresponding set of objects, compared to children who played a control game (Sella, Tressoldi, Lucangeli, & Zorzi, 2016). Promising results have also been shown in other computer-based games. For example, kindergarteners played one of two adaptive computer games for eight 25-min sessions; one game focused on comparing the quantities of numbers, and the other focused on learning procedural and conceptual counting knowledge. Compared to a control group, children in both conditions improved on counting and arithmetic after the intervention, and the gains remained six months later. However, no improvements were found in children’s number line estimation performance (Toll & Van Luit, 2013a).

**Domain-General Training of Mathematical Achievement**

Based on Geary and Hoard’s (2005) model, underlying cognitive systems play a critical role in mathematical development as well and should also be targeted for intervention. There is accumulating evidence for the efficacy of WM interventions for children in that there are generalizing effects to school-relevant measures (for recent reviews, see Diamond & Lee, 2011; Jaeggi & Buschkuehl, 2012; Wass, Scerif, & Johnson, 2012). Several studies have demonstrated that training on WM skills leads to improvements in school-age children’s and adults’ reading or mathematical achievement (Holmes & Gathercole, 2014; Karbach, Strobach, & Schubert, 2015; Loosli, Buschkuehl, Perrig, & Jaeggi, 2012; Witt, 2011). Although there are ongoing discussions concerning the replicability and the underlying mechanisms of observed improvements (Jaeggi, Buschkuehl, Jonides, & Shah, 2012; Shipstead, Redick, & Engle, 2012), the results of several recent meta-analyses are encouraging in that they do show generalizing effects beyond the trained domain (Au et al., 2015, 2016; Karbach & Verhaeghen, 2014; Wass et al., 2012; Weicker, Villringer, & Thöne-Otto, 2016, but see Melby-Lervåg & Hulme, 2013).
Existing interventions, however, have predominantly targeted older children, thus, we have only minimal knowledge concerning the effects of WM training in kindergartners, especially from low-income backgrounds. For example, Kroesbergen, van ’t Noordende, and Kolkman (2014) found that training kindergarten children’s WM skills improved their counting skills, but they found minimal improvement in their non-symbolic numerical magnitude knowledge. Understanding how WM training can promote low-income children’s symbolic numerical knowledge is particularly important since this is where the greatest SES-related disparities are present (Jordan & Levine, 2009). Based on the assumption that plasticity is even greater at a younger age (Garlick, 2002; Wass et al., 2012), it is likely that improving WM would be beneficial for this population with the potential to support early numeracy skills (Diamond & Lee, 2011).

Individual Differences and Training Outcomes

An important growing trend in intervention research is to understand for whom and under what conditions training can improve children’s outcomes. One factor that likely influences the benefits of interventions is participants’ initial abilities. For example, playing linear board games is more effective for preschoolers with lower initial numerical knowledge (Ramani & Siegler, 2011; Siegler & Ramani, 2009). Similarly, pre-existing individual differences in WM ability seem to be related to training success in that those with lower ability tend to improve more (Au et al., 2015; Jaeggi, Buschkuehl, Jonides, & Shah, 2011). One possibility is that lower initial abilities allow for more growth from training. However, other studies have found that initial mathematical knowledge is positively correlated with increases in mathematical achievement test performance (Geary, 2006).

Another factor that seems to be important is children’s engagement in activities. Number board games when played in a small group can keep children engaged during a play session (Ramani, Siegler, & Hitti, 2012). Similarly, for domain-general training in WM, engagement is related to children’s gains on transfer measures (Jaeggi et al., 2011; Jaeggi, Buschkuehl, Shah, & Jonides, 2014). WM training activities can be challenging, therefore, maintaining engagement during the activities is essential for benefiting from the activities. Understanding how individual differences influence children’s gains can be critical for implementing interventions widely.

The Current Study

The present study had three main goals. The first was to examine whether interventions targeting either numerical knowledge (domain-specific skills) or WM (domain-general skills) improved kindergarten children’s numerical knowledge. Both interventions were tablet games that were adapted from previous theoretical and empirical work. The numerical knowledge training game was a 0-100 number board game linearly-arranged in a 10 x 10 array adapted from Laski and Siegler (2014), who found that the numerical magnitude knowledge of low-income kindergarten children improved from playing the game one-on-one with an experimenter for six sessions across three weeks. The authors propose the design of the semi-linear board game clearly displays the base-10 system (i.e., the row represent the decades and the columns represent the units), which could be critical for children to learn about the numerical magnitudes of larger numbers. The WM training game was adapted from previous research that successfully trained transfer measures in elementary-school aged children and older adults (Buschkuehl et al., 2008; Loosli et al., 2012). Based on Geary and Hoard’s (2005) model, we hypothesized that both the domain-specific and domain-general game would improve children’s numerical knowledge more than children who did not play the games.
The second goal was to examine whether the interventions also improved children’s WM skills. There is a large literature demonstrating an association between children’s WM skills and their mathematics performance, however, since these relations are primarily correlational in nature, the direction and natures of these relations so far is unknown. Domain-general and domain-specific skills likely have a reciprocal relationship with greater developing skills in one can advance skills in the other. Therefore, we hypothesized that both training games would improve children’s WM compared to a control condition.

The third goal was to examine how individual differences predicted children’s outcomes. We examined whether pre-existing ability (i.e. children’s existing numerical magnitude knowledge) predicted their learning from the training. We also tested whether children’s engagement during the training activities predicted children’s outcomes on the numerical knowledge and WM measures. Self-reports of engagement have been important predictors of older children’s and adults’ learning from training (Jaeggi et al., 2011; Jaeggi et al., 2014), however, it is unclear whether kindergarteners would be able to give reliable self-reports that predict their learning.

**Method**

**Participants**

Participants were 81 kindergarteners, ranging in age from 5 years 4 months to 7 years 3 months ($M_{\text{age}} = 6$ years 0 months; 56% boys). Children were recruited from five classrooms in two public schools in a mid-Atlantic metropolitan area, which primarily serve families from lower-income backgrounds with high rates of free and reduced lunches (79% and 69% in each school respectively), a major indicator of poverty (Ingersoll, 1996). Demographic data are presented in Table 1. The race and ethnicity of the sample was diverse and parental education varied widely. Over 50% of the sample families made less than $45,000 a year. Two additional participants were recruited but left the school before the completing of the study. All classroom activities were conducted in English and used the same mathematics curriculum based on the Common Core State Standards. The focus of instructional time was to be on two areas: numbers (e.g., using written numbers to represent quantities and solve problems, counting objects in sets and representing and comparing numbers) and geometry (e.g., describing 2-dimensional shapes and 3-dimentional shapes, describing spatial relations and using spatial reasoning).

**Overview of Procedure**

The study consisted of 14 sessions conducted over the course of one to two months in the children’s classrooms. During sessions 1-2 and 13-14, an experimenter assessed children’s numerical knowledge and WM with three measures for each construct. Children were randomly assigned within classrooms to one of three conditions ($n = 27$ per condition): the domain-general training condition, the domain-specific training condition, or the control (business-as-usual) condition. During sessions 3 through 12, children in the two training conditions played videogame-like activities on tablets (Samsung Galaxy Tab 2 with a screen size of 10.1 in or 25.7 cm, measured diagonally) while supervised by an experimenter. Children played the game in the classrooms on an individual tablet while wearing headphones. Each session lasted approximately 15-20 min.
Table 1
Descriptive Statistics of Demographic Variables

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants (n = 81)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in months, M (SD)</td>
<td>72.11 (4.31)</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>44</td>
</tr>
<tr>
<td>Male</td>
<td>56</td>
</tr>
<tr>
<td>Race and Ethnicity of Children (%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>46</td>
</tr>
<tr>
<td>Caucasian/White</td>
<td>18</td>
</tr>
<tr>
<td>Mixed Race</td>
<td>12</td>
</tr>
<tr>
<td>African American or Black</td>
<td>11</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>11</td>
</tr>
<tr>
<td>Maternal Education (%)</td>
<td></td>
</tr>
<tr>
<td>Some high school</td>
<td>11</td>
</tr>
<tr>
<td>High school diploma/GED</td>
<td>32</td>
</tr>
<tr>
<td>Some college/vocational training or a 2-year college degree</td>
<td>21</td>
</tr>
<tr>
<td>College degree or post-graduate degree</td>
<td>30</td>
</tr>
<tr>
<td>Unreported</td>
<td>6</td>
</tr>
<tr>
<td>Paternal Education (%)</td>
<td></td>
</tr>
<tr>
<td>Some high school</td>
<td>16</td>
</tr>
<tr>
<td>High school diploma/GED</td>
<td>30</td>
</tr>
<tr>
<td>Some college/vocational training or a 2-year college degree</td>
<td>25</td>
</tr>
<tr>
<td>College degree or post-graduate degree</td>
<td>23</td>
</tr>
<tr>
<td>Unreported</td>
<td>6</td>
</tr>
<tr>
<td>Household Income (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Less than $15,000</td>
<td>12</td>
</tr>
<tr>
<td>$15,000 - $30,000</td>
<td>28</td>
</tr>
<tr>
<td>$31,000 - $45,000</td>
<td>16</td>
</tr>
<tr>
<td>$46,000 - $59,000</td>
<td>11</td>
</tr>
<tr>
<td>$60,000 - $75,000</td>
<td>10</td>
</tr>
<tr>
<td>$76,000 - $100,000</td>
<td>5</td>
</tr>
<tr>
<td>$101,000 or greater</td>
<td>9</td>
</tr>
<tr>
<td>Household Languages (%)</td>
<td></td>
</tr>
<tr>
<td>Monolingual English speakers</td>
<td>67</td>
</tr>
<tr>
<td>Exposed to more than one language</td>
<td>33</td>
</tr>
</tbody>
</table>

<sup>a</sup>Race and ethnicity were unreported for 2% of the children. <sup>b</sup>Household income was unreported for 9% of the families.

Conditions

The games used in both training conditions were designed to be engaging by incorporating video game-like features and artistic graphics (Gee, 2007; Malone & Lepper, 1987; Prensky, 2001). For example, the children earned gold coins for correct answers in both games. The games in both conditions also included 8 themes with the same characters that were tied to a background story (e.g., a magical pond, a pirate ship, outer space, or coral reef). All children played the same theme for each of the first 8 sessions (one per session), and then picked the theme of their choice for the remaining two training sessions.
Domain-Specific Condition: Numerical Knowledge Training

In this tablet game, children were presented with a 10 x 10 matrix numbered 1-100 with numbers in each row increasing from left to right (Figure 1a). Children were given a game token and were told that their opponent was the computer. For each turn, children spun by tapping the electronic spinner containing the numbers 1-6 pictured next to the game board. Children moved their token the number on the spinner by touching each square on the board game on the way to the final square and by pushing the “DONE” button, indicating when they were finished moving. Then the computer’s token moved. The computer would say the number in each space as the token moved. The computer also provided feedback if the children moved too many or too few spaces. Children were able to see the computer’s token during the game. Once every three turns, the game pieces were removed from the game board, and the screen showed the numbers the two tokens were on and children were prompted to indicate which character was leading by touching the number (Figure 1b). This game provided children with a visual cue to numerical magnitudes, as well as practice comparing the magnitudes of numbers. The dependent variables were the number of times feedback was given in response to an error (i.e., the number of turns children made an error) and the proportion of the magnitude comparison questions answered correctly.

![Figure 1a](image1a.png) ![Figure 1b](image1b.png)

Figure 1. Domain-specific condition: Numerical magnitude knowledge training a) Example of a turn; b) Example of a character magnitude comparison question.

Domain-General Condition: Working Memory Training

This tablet game was modeled after previous work (Buschkuehl et al., 2008; Loosli et al., 2012). Children were presented with a sequence of identical characters that varied only in color (Figure 2). Some of the characters were presented upside-down, and some were presented right-side up. Children had to indicate the orientation of each character by touching one of the corresponding buttons in the bottom corners of the screen. In the second part of the activity, children had to recall the sequence of colors of the characters presented in the first part by clicking on the appropriately colored characters from the display in the correct order. At the end of each trial, children received feedback on whether the recall was correct or not. The initial and lowest number of characters that was presented was two, but set size increased until children were not able to recall the sequence accurately anymore. If children were unsuccessful on recalling a sequence, the set size was decreased by one on the next trial, and it stayed the same if the children made an error on the orientation decision (or took longer than 3 seconds to make the orientation decision). Thus, the task was responsive to individual performance and therefore remained challenging, but did not overwhelm the children. The game also came in the same eight themes as described previously. The dependent measures were the average and the maximum size of the set of characters recalled per session and across all 10 sessions, as well as the accuracy and reaction times for the recall part of the activity.
Measures of Child Engagement and Motivation

After each training session, children answered three questions to assess their enjoyment, self-efficacy, and effort during the session. These questions have been validated in previous work (Jaeggi et al., 2011; McAuley, Duncan, & Tammen, 1989). Each question was printed at the top of the screen and read aloud by the computer. Response options were represented as faces ranging from a smile to a frown (5-point Likert scale). Answers to the questions were also read aloud by the computer with each face lighting up as the response was read. Children were instructed to press the “Done” button when they had chosen their response.

The first question was “How much did you enjoy this game?” with responses ranging from 1 = I really enjoyed the game to 5 = I really did not enjoy the game (Cronbach’s alpha across the 10 sessions = .81). The second question was “How good do you think you were at this game?” with responses ranging from 1 = I was really good at this game to 5 = I really was not good at the game (Cronbach’s alpha = .81). The third question was, “How hard did you try when you were playing this game?” responses ranged 1 = I tried really hard when I played the game to 5 = I really did not try hard when I played the game (Cronbach’s alpha = .84). The dependent variable was the average for each of the questions across the training sessions.

Control Condition

Children assigned to this condition participated in their standard kindergarten curriculum and only completed pre- and post-testing. Children were given the opportunity to play the tablet games for a session after completing the post-testing.

Outcome Measures

Children were given three assessments of numerical knowledge and three assessments of WM during the pretest and posttest sessions. The measures were administered in the same order.

Mathematical Knowledge

Number Line Estimation — To assess children’s numerical magnitude knowledge, children were administered a 0-100 number line estimation task. Children were presented with 20 cm lines on a tablet computer one at a
time. On each line, there was a “0” just below the left end of the line, and a “100” just below the right end. A number between 0 and 100 was displayed approximately 4 cm above the center of the line. The experimenter told the children that they would be playing a game in which they needed to mark the location of a number on a line. On each trial, the experimenter asked, “If this is where 0 goes (pointing) and this is where 100 goes (pointing), where does N go?” There were 26 trials of the following numbers 3, 4, 6, 8, 12, 14, 17, 18, 21, 24, 25, 29, 33, 39, 42, 48, 52, 57, 61, 64, 72, 79, 81, 84, 90, and 96 (Booth & Siegler, 2006). The numbers were presented in random order. The dependent variable was the accuracy of children’s estimates, measured by percentage of absolute error (PAE), which is computed using the formula: PAE = |estimate – estimated quantity|/scale of estimates) x 100 (Siegler & Booth, 2004).

**Number Identification** — Children were presented with 35 different Arabic numerals presented randomly one at a time on cards (adapted from Ramani & Siegler, 2008). The cards included all of the numbers between 1 and 19, and two numbers from each subsequent decade (20, 26, 31, 37, 41, 44, 53, 59, 62, 64, 75, 78, 83, 86, 92, 95). The dependent measure was the percentage of numerals correctly identified.

**Addition** — Children were asked to complete five simple arithmetic problems all with addends less than 10. The problems were written on paper, as well as presented orally by the experimenter. Children were encouraged to use their fingers or the paper to solve the problems. Two sets of problems were counterbalanced at pretest and posttest. The problems given across the two sets were 1+7, 2+3, 3+3, 5+2, 5+3, 1+4, 4+3, 5+4, and 6+3. The dependent measure was the total number of problems correctly solved.

**Working Memory**

**Forward Digit Span** — Children were read a list of numbers and asked to repeat the list in the same order presented (Wechsler, 1991). The number of items in the list ranged from a span of two and increased to a span of six. Four trials at each span level were given unless children correctly answered the first three trials in which case the last trial was skipped and the next level was administered. The task was discontinued if children incorrectly answered two of four trials at a particular span level. The dependent measure was the total number of trials that were correctly recalled.

**Backward Digit Span** — Similar to the forward digit span, children were read a list of numbers and asked to repeat the list in reverse order (Wechsler, 1991). Again, the number of items in the list ranged from a span of two and increased to a span of six, and each span level contained four trials unless children correctly answered the first three trials at a span level. The task was discontinued if children incorrectly answered two of four trials at a span level. The dependent measure was the total number of trials that were correctly recalled.

**Following Directions** — Children were seated in front of an array of familiar items (pencils, erasers, folders, boxes, rulers) in varying colors (blue, yellow, red) (Gathercole, Durling, Evans, Jeffcock, & Stone, 2008). On each trial, an experimenter asked the children to perform a spoken set of actions, such as “Pick up the blue ruler, and put it in the yellow folder.” The number of actions ranged from one and to seven (Levels 1-7), and each level contained four trials. Children completed all four trials on a given level unless they successfully answer the first three of four trials on a given level, in which case the last trial was skipped and the next level was administered. The number of actions was increased until children inaccurately answered two of four trials at a particular level. The dependent measure of span consisted of the total number of levels completed successfully plus partial credit for levels unsuccessfully completed. Specifically, if children completed the last
level with one correct answer, they received 1/3 point extra. If they answered two questions correctly on the last level, they received 2/3 point extra (cf. Holmes, Gathercole, & Dunning, 2009).

**Results**

**Preliminary Analyses**

Preliminary analyses revealed no age, gender, parent education, or classroom differences between groups (all $p_s > .38$). All of the children in both experimental conditions completed the 10 training sessions. However, there were missing data for the training measures, which were either due to experimenter or tablet error. In such cases, missing data was interpolated between the former and later session based on the individual training scores and measures of child engagement. This represented less than 5% of all of the tablet data.

**Training Measure**

For the number training game, the children played one game per session. Figure 3a presents the percentage of the character magnitude comparison questions (i.e., ‘who is leading?’) answered correctly during the games, which shows children were fairly accurate in their responses during the games. On average, children answered eight magnitude comparison questions per session and answered 83% ($SD = 12\%$) of those questions correctly. Figure 3b presents the number of turns feedback was given in response to a mistake, which shows a decline in the number of turns from the first session to the sixth session. After session 8, however, there is an increase in the number of error prompts received. On average, the children received corrective feedback for errors on 7.59 ($SD = 4.43$) occasions while playing the games.

![Figure 3a: Character Magnitude Comparisons](image1)
![Figure 3b: Feedback Prompts](image2)

*Figure 3. Training performance while playing the numerical magnitude game: a) the percentage of the magnitude comparison questions answered correctly (i.e., Who is leading?), b) the number of times feedback was given in response to an error.*

As for the WM game, the children completed 29 trials ($SD = 7.5$) on average per training session and reached an overall set size of 2.4 ($SD = .19$). Figure 4a illustrates the maximum set size for each training session for children who played the WM training game. The highest set size consisted of six items, which was only reached by one child; however, nineteen children (70%) reached a set size of 5 items. Figures 4b and 4c illustrate the reaction times as well as the accuracy during the recall part of the game, both of which significantly changed from the first to the last session (RT: $t(26) = 4.08$, $p < .001$, $d = .79$; accuracy: $t(26) = 3.07$, $p = .005$; $d = .59$).
Descriptive data as well as re-test reliabilities and effect sizes are reported in Table 2. We first tested whether the groups differed at pretest with a separate ANOVA for each outcome measure. No significant differences were found between groups on any measure at pretest.

**Outcome Measures**

We then investigated transfer by calculating two multivariate analyses of variance (MANOVA) with group (domain-general, domain-specific, or control) as the between factor and the differences between post- and pretest scores (from now on termed gain scores) as dependent variables using the three outcome measures for each domain separately (mathematics: number line PAE, numeral identification, and arithmetic performance; WM: following directions, forward and backward digit span).\(^1\) We report Pillai’s V as an F-statistic, assuming that it yields the most robust outcome. Further, in case of a significant group effect, we calculated univariate ANOVAs and used planned comparisons (Helmert contrasts) to follow up the ANOVAs in order to get at differential group effects, specifically, we ran the following two orthogonal contrasts: control group versus both experimental groups, and number game group versus WM game group\(^2\). Our central prediction was that there would be greater gains on the outcome measures from pretest to posttest for children in the two training conditions compared to the control condition. In contrast, we did not have specific predictions comparing the efficacy of the two intervention groups.

---


---

**Figure 4.** Training performance for the WM game: a) maximum set size recalled for each training session, b) average reaction time (RT) during the recall part of the game, c) average accuracy during the recall part of the game.
Table 2
Descriptive Data for the Transfer Measures as a Function of Group

<table>
<thead>
<tr>
<th>Group, Measure</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>Pre vs. Post</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
</tr>
<tr>
<td>Number Game Group (n = 27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number line (% Absolute Error)</td>
<td>23.82</td>
<td>7.71</td>
<td>7.85</td>
</tr>
<tr>
<td>Numeral Identification (% Correct)</td>
<td>84.66</td>
<td>17.79</td>
<td>22.86</td>
</tr>
<tr>
<td>Arithmetic (% Correct)</td>
<td>74.81</td>
<td>34.46</td>
<td>0.00</td>
</tr>
<tr>
<td>Working Memory Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Following instructions</td>
<td>2.33</td>
<td>0.53</td>
<td>1.33</td>
</tr>
<tr>
<td>Digit Span: Forward</td>
<td>8.52</td>
<td>1.70</td>
<td>5.00</td>
</tr>
<tr>
<td>Digit Span: Backward</td>
<td>2.85</td>
<td>1.63</td>
<td>0.00</td>
</tr>
<tr>
<td>Working Memory Game Group (n = 27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number line (% Absolute Error)</td>
<td>24.78</td>
<td>8.68</td>
<td>10.34</td>
</tr>
<tr>
<td>Numeral Identification (% Correct)</td>
<td>77.14</td>
<td>24.24</td>
<td>20.00</td>
</tr>
<tr>
<td>Arithmetic (% Correct)</td>
<td>76.30</td>
<td>32.83</td>
<td>0.00</td>
</tr>
<tr>
<td>Working Memory Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Following instructions</td>
<td>2.31</td>
<td>0.58</td>
<td>1.00</td>
</tr>
<tr>
<td>Digit Span: Forward</td>
<td>8.70</td>
<td>2.16</td>
<td>5.00</td>
</tr>
<tr>
<td>Digit Span: Backward</td>
<td>2.81</td>
<td>1.97</td>
<td>0.00</td>
</tr>
<tr>
<td>Control Group (n = 27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number line (% Absolute Error)</td>
<td>24.39</td>
<td>8.64</td>
<td>7.95</td>
</tr>
<tr>
<td>Numeral Identification (% Correct)</td>
<td>86.56</td>
<td>21.58</td>
<td>37.14</td>
</tr>
<tr>
<td>Arithmetic (% Correct)</td>
<td>80.00</td>
<td>31.87</td>
<td>0.00</td>
</tr>
<tr>
<td>Working Memory Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Following instructions</td>
<td>2.38</td>
<td>0.69</td>
<td>1.33</td>
</tr>
<tr>
<td>Digit Span: Forward</td>
<td>9.00</td>
<td>2.11</td>
<td>6.00</td>
</tr>
<tr>
<td>Digit Span: Backward</td>
<td>3.30</td>
<td>1.77</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note. r = re-test reliability (Pearson Correlation); ES = Effect size that accounts for the correlation between the pre- and posttest measures,

\[
ES = \frac{(\text{Mean}_{\text{Post}} - \text{Mean}_{\text{Pre}}) / \sqrt{SD^2_{\text{Pre}} + SD^2_{\text{Post}}} - 2r_{\text{PrePost}} * SD_{\text{Pre}} * SD_{\text{Post}}}}{}.
\]

*p < .05. **p < .01. ***p < .001.

Mathematics outcomes — The MANOVA using the three mathematics tasks as dependent variables was significant \(F(6, 154) = 2.40, p = .031, \eta^2 = .09\). Univariate ANOVAs revealed that the effect was mainly driven by the number line task \(F(2, 78) = 6.04, p = .004, \eta^2 = .13\), while the effects for the other two tasks were not significant (both \(p > .39, \eta^2 < .03\)). Planned contrasts for the number line task revealed that the two experimental groups significantly outperformed the control group in terms of gain from pretest to posttest \((t(78) = 2.60, p = .005 (1-tailed))\). Furthermore, children playing the number game showed larger gains than the children who played the WM game \((t(78) = 2.30, p = .02 (2-tailed))\). As illustrated in Figure 5, the effect sizes for the gains from pretest to posttest were substantial in the number game group and more than twice as large.
than those in the WM group ($d = 1.48$ vs. $d = .68$; both $p \leq .001$), but most importantly, both effect sizes were substantially larger than those of the control group, who did not significantly improve from pretest to posttest ($d = .28$, $p = .13$). Similarly for the number identification task, children who played both, the number game or the WM game significantly improved from pretest to posttest (number game group: $d = .68$; WM game group: $d = .72$), while children in the control group did not ($d = .32$). There were no significant improvements in the arithmetic task in any of the groups (all $d < .30$).

![Observed net effect sizes (Cohen's d) of performance gains from pre-test to post-test for the Math and WM outcome measures. Bars show net effect sizes (standardized changes in the trained group minus standardized changes in the control group), separately for the participants who completed the math intervention (dark bars) and those who completed the WM intervention (light bars).](image)

**Working memory outcomes** — The MANOVA using the three WM tasks as dependent variables was not significant ($F(6, 154) = .46$, $p = .84$, $\eta^2_p = .02$). Despite this lack of overall effects, it is important to note that children who played the number game significantly improved their performance in the backward digit span from pretest to posttest ($d = .43$, $p = .03$), while the improvement was only marginally significant in the WM group ($d = .35$, $p = .08$). However, the improvement in the control group was far from significant ($d = .15$, $p = .43$). There were no significant improvements in the two other WM measures in any of the groups (all $d < .15$).

**Child Engagement, Training Performance and Outcome Measures**

Next we examined whether children’s training performance and engagement ratings during the games related to children’s numerical knowledge and WM skills. First, we compared children’s engagement between the two training conditions using a Mann-Whitney U comparison for each of the measures (Table 3). The analyses indicated there were no differences between the two training games on any of the three questions of enjoyment, self-efficacy, and effort ($p > .42$).
Table 3
Means and Standard Deviations of the Self-Reported Engagement Measures for the Number and WM Training Games

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number Game</th>
<th>WM Game</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>2.02</td>
<td>.88</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>2.11</td>
<td>.90</td>
</tr>
<tr>
<td>Effort</td>
<td>1.98</td>
<td>.92</td>
</tr>
</tbody>
</table>

Note. Response options were represented as faces ranging from a smile to a frown (5-point Likert scale). On the scale, a value of 1 represents more enjoyment, self-efficacy, and effort, respectively. A value of 5 represents less enjoyment, self-efficacy, and effort, respectively.

Second, we examined whether children’s improvements in the two groups varied by children’s initial numerical magnitude knowledge. The correlations between initial numerical knowledge and improvement in the number line task were $r = -.52$, $p = .005$ (number game group) and $r = -.25$, $p = .21$ (WM game group). This suggests that for children playing the number game, the lower the initial magnitude knowledge (as indicated by higher PAE scores), the greater gains on the number line task, whereas this association was not as strong for children playing the WM game.

To further illustrate this point, we conducted a median split of all of the children based on their pretest performance on the number line task (median PAE score = 24%, range = 8% to 44%). Paired samples t-tests of children in the number condition showed that both children with lower and higher initial numerical knowledge improved from pretest to posttest. Specifically, children who played the number game with lower initial knowledge improved from their PAE from 29% to 21% ($t_{(15)} = 7.95$, $p < .001$, $d = 1.61$), and children with higher initial knowledge improved their PAE from 17% to 12% ($t_{(10)} = 3.58$, $p < .01$, $d = .79$). However, for children who played the WM game, only children with lower initial knowledge significantly improved from pretest to posttest with PAE scores decreasing from 32% to 28%, ($t_{(12)} = 3.69$, $p < .01$, $d = .59$), whereas there was only a minimal decrease for children with greater initial accuracy on the number line task, 18% to 16% PAE ($t_{(13)} = 1.62$, $p = .13$, $d = .39$). There was no change for the children with lower or higher initial knowledge in the control condition ($p > .20$).

Next we calculated correlations between the engagement measures, training measures, and each of the outcome measures at pretest and posttest separately for the two training conditions. Since lower scores for the number line estimation task indicate higher accuracy, we reversed this score for the subsequent analyses to be consistent with the other measures to ease interpretability. As shown in Table 4 for the number training group, only children’s rating of effort was correlated with performance on arithmetic at posttest, indicating higher ratings of effort during the game was associated with greater arithmetic performance on the posttest. Children’s training performance on the character magnitude comparison questions during the game and the number of times children received feedback during the game was correlated with two measures of numerical knowledge, numeral identification and number line estimation. Feedback was also related to children’s performance on the backward digit span at pretest and addition performance at posttest.
Table 4

Correlations Between Child Engagement, Training Performance, and Outcome Measures for the Number Game Condition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number Line PAE</td>
<td>No Id</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>-.08</td>
<td>-.11</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>-.31</td>
<td>-.34</td>
</tr>
<tr>
<td>Effort</td>
<td>.12</td>
<td>-.18</td>
</tr>
<tr>
<td>Character Mag Compare</td>
<td>.48*</td>
<td>.63**</td>
</tr>
<tr>
<td>Feedback</td>
<td>-.49**</td>
<td>-.59**</td>
</tr>
</tbody>
</table>

Note. PAE = Percent absolute error; Character Mag Compare = Character Magnitude Comparison; No Id = numerical identification.

*Percent absolute error (PAE) scores from the number line task are reversed to be consistent with the other measures with higher scores indicating greater accuracy.

Table 5

Correlations Between Child Engagement, Training Performance, and Outcome Measures for the WM Training Condition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number Line PAE</td>
<td>No Id</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>-.14</td>
<td>-.06</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>-.21</td>
<td>-.05</td>
</tr>
<tr>
<td>Effort</td>
<td>-.27</td>
<td>-.09</td>
</tr>
<tr>
<td>Average Set Size</td>
<td>-.02</td>
<td>.22</td>
</tr>
<tr>
<td>Max Set Size</td>
<td>.11</td>
<td>.37</td>
</tr>
</tbody>
</table>

Note. PAE = Percent absolute error; No Id = numerical identification.

*Percent absolute error (PAE) scores from the number line task are reversed to be consistent with the other measures with higher scores indicating greater accuracy.

*p < .05. **p < .01. ***p < .001.
As shown in Table 5, for children in the WM training group, the engagement measures were correlated with the outcome measures at pretest and posttest. For example, children’s responses on the effort questions were related to children’s performance on the number line estimation and addition performance at posttest. Children’s enjoyment on the game was related to the forward digit span at posttest. In regards to training performance, average set size and maximum set size was correlated with children’s numeral identification task at posttest. Average set size was also correlated with numeral identification.

Overall, there were stronger associations between the training measures for children who played the number game than for children who played the WM game. However, child engagement was more strongly correlated with the outcome measures in those children who played the WM game as compared with those who played the number game.

Finally, to examine whether measures of engagement and training performance predicted children’s posttest performance on the number line estimation, we conducted two hierarchical linear regression analyses. Regressions were only conducted for the number line estimation task since this was the only measure showing significant group effects and yielded the largest improvements for both training groups. Given the two interventions required different amounts of engagement, we conducted the regressions separately for the two training groups. For both analyses, we first controlled for pretest scores, then entered the three measures of engagement as a block followed by the two measures of training performance for each of the two training groups (i.e., magnitude comparison and feedback for the number training group; average set size and maximum set size for the WM training group).

As shown in Table 6, for the number game group, in Model 3 pretest number line performance ($p < .001$) was a significant predictor for children’s number line estimation accuracy at posttest; however, neither ratings of engagement nor training performance were significant predictors, after controlling for their pretest performance. Model 3 accounted for 81% of the variation in children’s posttest number line estimation accuracy.

Table 6

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Parameter estimate (standardized)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td>Pretest Number Line PAE</td>
<td>.85***</td>
</tr>
<tr>
<td>Engagement Measures</td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>.31</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>.00</td>
</tr>
<tr>
<td>Effort</td>
<td>-.14</td>
</tr>
<tr>
<td>Training Performance</td>
<td></td>
</tr>
<tr>
<td>Character Magnitude Comparison</td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>.08</td>
</tr>
<tr>
<td>$R^2$ statistic (%)</td>
<td>73</td>
</tr>
<tr>
<td>$F$ statistic</td>
<td>66.94***</td>
</tr>
<tr>
<td>Degrees of freedom for $F$ statistic</td>
<td>1, 25</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
As shown in Table 7, similar regressions conducted for the WM condition (Model 3) revealed that self-reported effort ratings significantly predicted posttest number line estimation ($p < .05$) after controlling for pretest number line estimation performance ($p < .001$). These variables combined explained 84% of the variation in children's number line estimation posttest performance.

### Discussion

This study tested whether playing tablet games that focused on domain-specific and domain-general training is an efficacious means for improving the numerical knowledge and WM of kindergarten children primarily from lower-income backgrounds. Playing both the number game and the WM game for 10 sessions on the computer tablet led to improvements in children’s numerical knowledge, particularly in the area of numerical magnitude understanding. These improvements were greatest for children who had lower pre-existing numerical magnitude knowledge, and furthermore, their self-reported effort predicted the learning outcome.

### Impact of Playing Domain-Specific and Domain-General Training Games

Our study provides evidence that both domain-specific knowledge and domain-general processes are critical for children’s numerical knowledge. Our study has tested an established conceptual framework of mathematics development (Geary & Hoard, 2005), and provides experimental evidence to support it. Previous studies have typically examined the relative contributions of numerical magnitude knowledge and WM using correlational designs (Alloway & Passolunghi, 2011; Geary, 2011; McKenzie, Bull, & Gray, 2003). The present findings go beyond those correlational approaches and demonstrate that providing children with training in domain-specific knowledge and domain-general skills can enhance children’s numerical magnitude knowledge. To our knowledge, the present work is the first to test the benefits of numerical magnitude knowledge and WM training within the same experiment by focusing on children from primarily low-income backgrounds. The present study...
was important to implement with kindergarteners since this first year of formal schooling is critical for laying foundational mathematics knowledge as well as WM abilities, because deficits in either can have consequences for mathematical achievement (Geary, 2006; Toll & Van Luit, 2013b). Previous studies training low-income or at-risk kindergarteners’ numerical magnitude knowledge have not included a training component for WM (Laski & Siegler, 2014; Praet & Desoete, 2014; Toll & Van Luit, 2013a). Additionally, WM training studies have primarily focused on older children and adults and have not focused on low-income populations (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi et al., 2011; Jaeggi et al., 2010; Loosli et al., 2012).

Several of the findings from the numerical magnitude and WM training games used in the present study replicate previous research in important ways. Specifically, consistent with Laski and Siegler (2014), we found that playing a 0-100 number board game improved kindergarten children’s number line estimation accuracy and numeral identification skills. These results are also consistent with studies with younger preschoolers that have shown playing a 0-10 linear number board game improved Head Start children’s number line estimation and numeral identification (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009). Our findings for the domain-general training are also consistent with the few studies that have examined WM training in young children, which have similarly shown that such both computerized and paper-and-pencil training programs can benefit their mathematical knowledge (Kroesbergen et al., 2014; Passolunghi & Costa, 2016).

However, it is also important to note that the findings of the present study extend this previous research in several important ways. First, consistent with the cognitive alignment framework (Laski & Siegler, 2014), our findings show that a number board game with a theoretically motivated design can be successfully translated into a tablet game. Although board games can have numerous advantages, such as social engagement and social learning from parents, teachers, or peers, there are also disadvantages, namely, board games require play partners to be available. The tablet game provides an alternative opportunity for children to play the board game when others may not be available. To further understand the role of social interactions during game play, future research could compare the effects of playing a computer version of the board game versus playing the physical version of the game with an adult.

Second, our findings show that tablets are a promising technological tool that could be utilized to improve children’s numerical knowledge. Although previous studies have found that non-tablet computer games targeting either mathematical skills (Räsänen et al., 2009; Wilson et al., 2009) or WM (Passolunghi & Costa, 2016) can improve young children’s numerical knowledge, to our knowledge, this is the first study to show similar improvements using tablet computers. This is significant because technological tools, such as iPads, have considerable educational potential, since they are more portable and affordable than laptops (Leoni, 2010). Further, interacting with electronic tablets is very intuitive and largely self-explanatory; kindergartners did not have technical issues while playing the games, because learning how to use a mouse was unnecessary.

Third, our findings provide evidence that experimentally tested tablet games can produce improvements in young children’s numerical knowledge. Parents and teachers of young children appear to value the use of technology, and specifically tablets, such as iPads, to promote learning (Verenikina & Kervin, 2011). However, despite the numerous apps available for young children’s mathematics learning, very few of them are theoretically grounded or tested (Hirsh-Pasek et al., 2015). The games used in the present study represent beneficial tablet game options to parents and teachers to use in various settings.
Individual Differences in Influence of Training Games

The present study adds to a growing area of intervention research which seeks to better understand for whom and under what conditions training activities are most beneficial to children’s learning. We found that self-reports of greater effort put into the WM game predicted children’s number line estimation accuracy after playing the games, even after controlling for their initial number line performance. This suggests that level of engagement during the WM game played a critical role in its influence on children’s number line estimation performance. These findings are consistent with previous WM training that has shown that children’s self-perceptions about the games are related to their improvements during the game. Specifically, children who felt the WM training game was challenging, but not overwhelming improved more from the training. Further, greater gains in the WM training game were associated with improvements in transfer tasks (Jaeggi et al., 2011).

We also examined how initial ability impacted children’s improvements in number line estimation. We found that playing the number game improved numerical magnitude knowledge for children with lower initial ability as well as for children with higher initial ability. In contrast, only children with lower initial numerical magnitude knowledge improved from playing the WM training game. It is likely that higher numerical magnitude knowledge indicates that children have a higher skill level and therefore they rely less on their WM. However, children with lower initial ability likely rely more on their domain-general skills to do those (non-automatic) tasks, then improving WM might be the right approach. This suggests that for children with more advanced numerical knowledge providing greater training with games that target their domain-specific skills may be more beneficial than training that targets their domain-general skills.

Limitations and Conclusions

There are several limitations of this present study. First, although we observed improvements in children’s numerical knowledge, we did not find improvements in children’s arithmetic performance and only limited improvements in WM. For the arithmetic problems this was likely due to ceiling effects with 73% of the children answering 4 or 5 of the problems correctly at pretest. For the WM measures, it is likely that these measures relied too heavily on children’s language skills, making the tasks difficult given that 33% of parents in our diverse sample reported that their children are exposed to more than one language at home. Thus, this reliance on language skills likely limited the reliability of the measures (see Table 2), which may have accounted for the fact that there was far transfer to the mathematics tasks from playing the WM tablet game, but not near transfer to the WM tasks. Future research that includes less verbal or non-verbal measures of the WM would be important to incorporate, especially for young children from diverse backgrounds.

Second, the training data indicate that children only minimally improved their performance in the training tasks (Figures 3 and 4). Specifically for the WM game, even though the children became significantly faster and more accurate indicating that they became more efficient across the intervention period, the amount of items they were able to hold in memory (set size) remained largely the same. It may be that that these young children were already at or close to their maximum performance, which restricted their room to improve. Alternatively, it is possible that the training sessions occurred too far apart from one another, that is, children typically only played the games twice a week to reduce boredom. However, despite the known benefits of spacing effects (Wang, Zhou, & Shah, 2014), spacing sessions too far apart might make it too difficult for children to build upon...
previous performance. Overall, given that training quality likely affects transfer (e.g. Jaeggi et al., 2011), the lack of training improvement might be another reason for the minimal improvement observed in the WM outcome measures. Still, the change in efficiency as expressed by their improved accuracy and RT scores might reflect a domain-general process that could have been the driving factor for the transfer effects to numerical knowledge.

Third, our control condition received their standard classroom instruction during the study. It is possible that playing a tablet game regardless of the content could have improved children’s numerical knowledge; however, the differential effects and individual differences analyses provide some evidence that the effects are likely tied to the specific tasks and not due to placebo effects (Au, Buschkuehl, Duncan, & Jaeggi, 2016). Nonetheless, studies that incorporate a control condition in which children play a tablet game without numerical content or WM demands will be important in order to further disentangle the specificity of the effect.

To conclude, this study replicates and extends previous research showing that playing number board games can improve children’s numerical magnitude knowledge. Further, our data provide some evidence that playing games that require domain-general skills (i.e., WM) can improve children’s numerical magnitude knowledge as well. Finally, these games can be implemented and played on a rapidly growing and easy-to-use technology, tablets. Future research has to demonstrate whether combining the two training approaches will result in enhanced effects compared to the outcome of each training approach on its own. Individual differences should be considered in future interventions, including assessments of sustained attention and engagement (particularly for WM games).

The changing technological landscape offers fresh opportunities to capitalize on popular, intuitive platforms such as the tablet computer (Pew Research Center, 2013) for improving young children’s numerical magnitude knowledge. Parents and teachers alike value the use of tablets to promote learning for young children (Verenikina & Kervin, 2011), and increased portability and affordability make tablets promising technological tools with considerable educational potential (Leoni, 2010), especially for young children who have not yet mastered the motor skills necessary to use a traditional personal computer. Thus, tablet computer games are an engaging, accessible medium through which to help close the early achievement gap in early mathematics understanding between children from lower-income backgrounds and children from higher-income backgrounds and to improve young children’s numerical knowledge in order to ensure a solid foundation for later mathematical understanding through adulthood.

Notes

i) The results of the outcome measures did not change when age was included as a co-variate. The MANCOVA for the three mathematics task remained significant for condition, $F(6, 152) = 2.20, p = .046, \eta^2_p = .08$), but was not significant for age, $F(3, 75) = .70, p = .554, \eta^2_p = .03$. The MANCOVA for the three working memory outcomes was neither significant for condition, $F(6, 152) = .46, p = .839, \eta^2_p = .02$, nor for age, $F(3, 75) = 1.27, p = .290, \eta^2_p = .05$.

ii) When univariate ANCOVAs with pretest scores as covariate were conducted to follow-up the MANOVA, the results are consistent to the results presented. Specifically, there were significant group differences on the number line estimation task, $(F(2,77) = 7.13, p = .001, \eta^2_p = .16)$, after controlling for pretest scores $(F(1,77) = 171.53, p < .001, \eta^2_p = .69)$. Using pre-test as covariate also does not change the pattern of results for the remaining effects.
Funding
This research was supported by a grant from the University of Maryland NSF-ADVANCE Interdisciplinary and Engaged Research Seed Grant Program.

Competing Interests
Martin Buschkuehl is employed at the MIND Research Institute, whose interest is related to this work and Susanne M. Jaeggi has an indirect financial interest in MIND Research Institute.

Acknowledgments
We would like to thank the administrators, parents, and children who supported and participated in this research.

References


Domain-Specific and Domain-General Training


