Author(s): Salminen, Jonna; Koponen, Tuire; Räsänen, Pekka; Aro, Mikko

Title: Preventive Support for Kindergarteners Most At-Risk for Mathematics Difficulties: Computer-Assisted Intervention

Year: 2015

Version:

Preventive Support for Kindergarteners Most At-Risk for Mathematics Difficulties: Computer-Assisted Intervention

Jonna Salminen and Tuire Koponen

Pekka Räsänen and Mikko Aro

Jonna Salminen, Department of Education, Special Education, University of Jyväskylä; Tuire Koponen, Niilo Mäki Institute, Jyväskylä; Pekka Räsänen, Niilo Mäki Institute, Jyväskylä; Mikko Aro, Department of Education, Special Education, University of Jyväskylä.

Correspondence concerning this article should be addressed to Jonna Salminen, Department of Education, Special Education, University of Jyväskylä, PO box 35 40014 Jyväskylä, Finland, jonna.salminen@jyu.fi

Preventive Support for Kindergarteners Most At-Risk for Mathematics Difficulties: Computer-Assisted Intervention

Weaknesses in early number skills have been found to be a risk factor for later difficulties in mathematical performance. Nevertheless, only a few intervention studies with young children have been published. In this study, the responsiveness to early support in kindergarteners with most severe difficulties was examined with two different computer programs. Two intervention groups were matched by age, visuo-spatial and phonological working memory, as well as early number skills. After a short and intensive computerized intervention, the results indicated significant intervention effects for verbal counting Wilcoxon *ES (r)* = .46, and dot counting fluency, *r* = .52, when practiced with GraphoGame Math, as well as for basic arithmetic, *r* = .63, when practiced with Number Race. The findings suggest that a targeted computerized practice can produce specific training effects in kindergarteners most at-risk for mathematics difficulties. The results are discussed with regard to practical implications for educational game development.

Mathematics difficulties (MD), typically seen as deficits in basic arithmetic (Geary, 2011), can be predicted rather reliably from kindergarten (e.g., McClelland, Acock, & Morrison, 2006; Morgan, Farkas, & Wu, 2009; Murphy, Mazzocco, Hanich, & Early, 2007). It is known that the difference in early number skills is notable between low-performing children and typically performing ones, and the gap seems to grow during the following years (e.g., Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Desoete & Grégoire, 2006; Jordan, Kaplan, Locuniak, & Ramineni, 2007). According to Geary (2011) the children with the most severe difficulties (scoring below the 10th percentile on a standardized mathematic achievement test) differ from their low achieving age-peers (scoring between the 11th and 25th percentile) as having even more persistent deficits (see also Chong & Siegel, 2008; Geary, 2013; Geary, Hoard, & Bailey, 2012). Such deficits or delays can be seen in processing numbers, learning arithmetic procedures, retrieving arithmetic facts, and in working memory (Geary, 2011). The growth rates of these sub-groups (performance below 10th, or between 11th and 25th percentiles) differ from each other when proceeding from kindergarten to later, primary grade levels (e.g., Geary, Hoard, Nugent, & Byrd-Craven, 2008; Morgan et al., 2009; Murphy et al., 2007). The children with the most severe difficulties fall even more behind. These findings underline the need to develop and investigate tools to effectively support such kindergarteners, especially those most at-risk for severe disability in learning arithmetic which per se is a critical feature of mathematics difficulties (Butterworth, Varma, & Laurillard, 2011; Geary, 2011). In order to improve such tools, it is important to understand what the key early number skills are, and how do they develop.

EARLY NUMBER SKILLS

It has been observed that young children can recognize small quantities by subitizing (Dehaene, 2011; Starkey & Cooper, 1980). This term describes the process by which children quickly see how many items are in a given space. Researchers have noted that this ability is inborn and is restricted to small quantities (e.g., Piazza, Mechelli, Butterworth, & Price, 2002; Reeve, Reynolds,

Humberstone, & Butterworth, 2012). The ability to discriminate small quantities is a prerequisite for learning the meaning of small number words such as *one*, *two*, and *three*, which are typically learned by children by the time they are 3.5 years old (e.g., Wynn, 1990). Enumerating quantities of more than three requires counting (see Dehaene, 2011, p. 56−59), or detecting the number of subsets for forming the whole quantity that is conceptual subitizing (Sarama & Clements, 2009).

The development of verbal counting involves several steps. Young children start verbal counting by repeating number words first as a rhyme, starting from number one, without necessarily knowing their numerical values (Fuson, 1988). After that, number words can be perceived as separate items and children learn the correct order of the small number words (Fuson, 1988). The development of verbal counting skill is important for later arithmetic performance at school age (e.g., Desoete & Grégoire, 2006; Koponen, Salmi, Eklund, & Aro, 2013; Lepola, Niemi, Kuikka, & Hannula, 2005). More specifically, it has been found that verbal counting predicts both procedural calculation and fact retrieval fluency (Koponen, Aunola, Ahonen, & Nurmi, 2007).

During early number skill development, the knowledge of number word sequence becomes integrated with the cardinality meaning of number words in order to be used in counting as a tool for determining the exact magnitude of objects, that is object counting. Hence, exact object counting demands understanding that the last counted number word represents the total amount of items, as well as other counting principles (e.g., one-to-one correspondence between count words and objects being counted, stable order of the count words and cardinality) (e.g., Clements, 2004; Sarama & Clements, 2009; Gelman & Gallistel, 1978; Wynn, 1992). It seems that object counting fluency remains rather stable between different skill-level groups (Reeve et al., 2012; 6 year old kindergarteners were followed for 5 years). Also, early number competence including both verbal counting and object counting as well as other numerical skills is found to highly predict later mathematics achievement at school (Jordan et al, 2007; Mazzocco & Thompson, 2005), and more specifically calculation fluency (Locuniak & Jordan, 2008) as well as applied problem solving

(Jordan, Kaplan, Ramineni, & Locuniak, 2009). Furthermore, when exact object counting skills develop, they can be used for composing and decomposing quantities (Baroody, 1987; Sarama & Clements, 2009) as well as for adding and taking away. These skills are prerequisites for fluency in later calculation skills. Respectively, non-fluency with arithmetic combinations is a critical characteristic of MD (Gersten, Jordan, & Flojo, 2005).

Children with atypical number skill development seem to have deficits in symbolic comparison, while the evidence for similar deficits in non-symbolic comparison remains rather contradictory (see Bartelet, Vaessen, Blomert, & Ansari, 2014; De Smedt, Noël, Gilmore, & Ansari, 2013). Kindergarteners with the weakest performance in counting and other early number skills (performance below 10%) appear to have a slower rate of growth than low performing children (performance from 11% to 25%) through the third grade, a trend that might continue in later grades (Murphy et al., 2007). Overall, the definitions of at-risk status (e.g., Mazzocco, 2005) and MD (e.g., Butterworth et al., 2011; Fuchs, 2005; Geary, 2004, 2011, 2013) are currently becoming more specific. Understanding the heterogeneity among the lowest performers has implications for early identification and suggests the need for intensified and individualized support (Geary, 2011).

COMPUTER-ASSISTED INTERVENTION

Recently, the use of computers in daily kindergarten activities has increased greatly. A variety of computers, laptops, and tablets are available with downward trend of costs. Researchers have considered the benefits of computer use for mathematics learning for decades. Computers can provide developmentally appropriate experiences for children (Clements, 2002), as well as motivate (Becker, 1992; De Smedt et al., 2013) and activate (Chambers & Sprecher, 1980) children. Furthermore, they provide immediate and continuous feedback as well as repetitive practice, all of which are found to be important for children with weak skills (Hasselbring, 1986). Li and Ma (2010) recently concluded that computer technology is more effective with regard to mathematics achievement in special needs students than in general education students. As Slavin and Lake

(2008) summarized, a reasonable use of computers can provide mathematics exercises tailored to individual needs, and adaptive software can identify child's strengths and weaknesses to fill possible gaps. These findings are consistent with other meta-analyses and reviews in which low achievers and at-risk learners progressed more than other students when computer-assisted intervention (CAI) was used (e.g., Kroesbergen & Van Luit, 2003; Kulik & Kulik, 1991; Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009). However, Räsänen (in press) reviewed that in recent decades the main trend of the effectiveness of CAI on numerical skills has been declining, not increasing. Similar observations on the trend have been made in three recent meta-analyses (Cheung & Slavin, 2013; Christmann & Badgett, 2003; Li & Ma, 2010).

The effectiveness of CAI has been difficult to establish in research literature due to varying study designs, target group definitions, and reports of the content being practiced (e.g., Räsänen et al., 2009; Seo & Bryant, 2009; Slavin & Lake, 2008). For this reason, concerning children in primary grades, Räsänen et al. (2009) were able to calculate effect sizes for only five CAI studies in which pre- and post-test scores with standard deviations for both intervention and control groups were reported. Seo and Bryant (2009) also faced several methodological problems in analyzing the effects of CAI studies in children with learning disabilities (see also Cheung & Slavin, 2013; Slavin & Lake, 2008).

There are only a few previous CAI studies of early number skills for young children (5−7-year olds) at-risk for difficulties in learning mathematics. Baroody, Eiland, Purpura, and Reid (2013) recently reported highly robust CAI effects for children at-risk of not learning the add-1 rule of basic addition. These children received the intervention in two stages: first, children played manual games, and then they had guided computer practice sessions. Researchers have conducted a similar study with significant effects for kindergarteners regarding the add-0 and add-1 rules (Baroody, Eiland, Purpura, & Reid, 2012). With the first graders at-risk for reading and math difficulties, the effect of CAI was significant for addition fact fluency (sum ≤ 18) (Fuchs et al.,

2006). Children's basic arithmetic (Praet & Desoete, 2014), numeral recognition (McCollister, Burts, Wright, & Hildreth, 1986), enumeration (Ortega-Tudela & Gómez-Ariza, 2006), and symbolic comparison (Wilson, Dehaene, Dubois, & Fayol, 2009) skills have also been significantly enhanced by CAI. Even in younger children (4-year olds) studies have shown positive effects on pre-mathematical knowledge (Howard, Watson, Brinkley, & Ingels-Young, 1994; Elliot & Hall, 1997). On the other hand, Din and Calao (2001) found no statistically significant results in mathematics when low SES children played several educational video games. Detailed information of these CAI intervention studies are summarized in Table 1.

Insert Table 1 here.

In the literature regarding intervention effectiveness, a central issue is the duration and intensity of the practice. In CAIs lasting for several months, a semester, or a whole year, the effects seem to be less clear than in shorter interventions of four weeks or less, irrespective of target group age (e.g., Kroesbergen & Van Luit, 2003; Kulik & Kulik, 1991). More focused interventions to specific skills, higher intensity, and more homogeneous target groups may explain the larger effect sizes in short interventions (Räsänen, in press). However, the finding also suggests that even very short but intensive interventions can be used to produce significant gains if the content of the practice is aligned with the needs of the learner.

THE CURRENT STUDY

This study examines the effects of two freely downloadable, adaptive mathematical computer programs on kindergarteners most at-risk for mathematics difficulties. Here, most at-risk status signifies poor performance in verbal counting ($\leq$ the 10th percentile), along with significantly slower dot counting, weaker basic arithmetic, as well as visuo-spatial and phonological working memory skills as compared to a reference group (not most at-risk for MD). In the current study, verbal counting level was used as inclusion criterion because it seems to be a strong predictor of later arithmetic achievement at school (Aunio & Niemivirta, 2010; Aunola et al., 2004; Koponen

et al., 2013; Lepola et al., 2005). Verbal counting was assessed with similar types of tasks as the aforementioned studies.

There were also other reasons for selecting verbal counting instead of all early number skills as inclusion criterion. We did not use object counting as inclusion criterion because it seems to be more effective at differentiating between the lowest and typically achieving children at somewhat older age levels (approximately 8−14 years old; e.g., De Smedt, et al., 2013). From a more technical point of view, the performance in our number comparison task was not used as inclusion criterion. The children were asked to determine, as quickly as possible, which of the two presented symbols was larger on the computer screen (on the left or right side on the screen) by clicking the respective mouse button. The time pressure increased the number of errors and even caused guessing in the most at-risk children. Thus, neither the fluency in number comparison nor the accuracy (not properly measured in the current study) were analyzed further, even though number comparison seems to differentiate children with and without (risk for) MD, and is related to arithmetic skills (e.g., Bartelet, Vaessen, et al., 2014; Skagerlund &Träff, 2014). Further, fluency in basic arithmetic at kindergarten age is related to familiarity with arithmetic symbols, as reflected in low general performance level at the pre-assessments in the current study. Therefore, basic arithmetic was not suitable as inclusion criterion.

The main purpose of this study was to examine if short and intensive practice with computer programs can support the early number skills (verbal counting, object counting or basic arithmetic) in kindergarteners most at-risk for MD. Here, short and intensive practice period means training for three weeks, 10−15 minutes per day (c.f., durations and intensiveness in earlier studies described in Table 1). We also examined the between-condition differences in potential intervention gain scores. Finally, we examined the association of the gain scores with the total intervention exposure. In this study, two intervention conditions were used. One focused on exact numerical processes, and the other on approximate numerical processes. Therefore, the total intervention exposure was

contrasted with the potential positive gains within conditions. Due to the fact that the evidence of CAI effectiveness in kindergarteners most at-risk for MD is still largely missing, exact hypotheses were not set for the current study.

METHOD

Participants

To conduct our study, we first obtained written permission from a municipal official in charge of day care in a city in Eastern Finland. Next, the official recruited voluntary teachers from day care centers to operate as coordinators for the intervention. We requested written permission from parents whose child took part in the kindergarten curriculum at any of these day care centers (*12* day care centers, altogether *236* kindergarteners). We informed parents of the purpose of the study, and of their right to discontinue the participation at any point. Of the resulting group of children, candidates ($n = 30$) were nominated into intervention group based on the teachers' observations of who needed extra support for early number skills. If the original number of children was 10 or less per kindergarten group, the teacher was asked to nominate one candidate. If the number of children was 11−21; 22−35; 36 or more per group, the teacher was asked to nominate 2; 3; and 4 candidates, respectively. To form the reference group ($n = 30$), teachers also nominated one peer-control for each candidate from the same kindergarten group. The peer-control was selected on the basis of having the nearest birthday to the candidate, and of not being in need of extra support for early number skills. Nomination was followed by two individual pre-tests of cognitive abilities, early number skills, and control measures for the candidates and peer-controls that were the reference children. All assessments were administered by the first author and a research assistant. Both have experience assessing young children, but were unfamiliar with the participants in this study. The sample was homogenous in cultural background, and all participants were native speakers of Finnish.

The findings concerning the total sample of children ($n = 60$) have earlier been published in a separate study (Räsänen et al., 2009). For the current study, only those target children who performed below the 10th percentile of the reference group (not most at-risk for MD) in verbal counting were included. Verbal counting (count on, count backward, and skip count by 2) was used as criterion task because of its importance in early number skill development, and in learning more complex mathematics at school age. Based on the inclusion criteria, the study sample consisted of 17 intervention children (7 % of the original sample; $n = 236$), and the reference group ($n = 30$; 13 boys, 17 girls; mean age = 78.8 months, $SD = 3.3$) was used only to determine the risk level in verbal counting, and for testing the test−retest reliabilities for early number skill measures. The children without the risk status for MD typically have mastered prerequisite numerical skills at kindergarten age. For example, in our study 23 of 30 children in the reference group reached the maximum score in counting skills at post-assessment (without extra support). For this reason, the reference group data are not analyzed further, and the group comparisons were not carried out between intervention and reference conditions.

In the current study, poor performance in verbal counting means that the majority of our intervention children ($n = 17$) could not even count correctly up to 20 (64.7 % of participants), and they also failed at more complex tasks, such as counting backward from 12 to 8 (76.5 %) or skip counting by 2 up to 10 (70.6 %) in February, during their kindergarten year. This skill level was comparable to the level of children scoring below the 10th percentile in a normative sample of Finnish kindergarteners ($n = 502$) that was collected for a nationally normed assessment test that included number knowledge, number concept, verbal counting, and non-verbal calculation tasks (see technical manual; Polet & Koponen, 2011). In this normative sample, collected in January-February during kindergarten year, 61.2 % of the poorest performers (lowest 10 %, $n = 49$) could not count on up to 20, 85.7 % could not count backward from 12 to 8, and 87.8 % could not skip count by 2 up to 10. Among the rest of the children (i.e. performance above the 10th percentile; $n =$

453) the corresponding failure percentages were 3.8 %; 15.9 %; and 21.2 %. In the reference group of the current study (not most at-risk for MD, $n = 30$) the respective percentages were 6.7 %; 16.7 %; and 16.7 %.

For ethical reasons, all parents provided written permission for their child to participate in both the assessments and the intervention. All parents were also informed of the research project as follows: The Ministry of Education and Culture in Finland (2007−2013) has funded a research project during which a research-based web service for learning challenges in early reading and mathematics will be created, and the effectiveness of certain educational computer games for early support will be studied. In addition, all children in the 12 participating day care centers were allowed to use the intervention programs after the actual study if their parents gave consent that they could do so.

Design and Materials

The study took place at day care centers for six weeks from February to April. During this period, all 17 participants followed the normal kindergarten curriculum. According to the Finnish National Board of Education (2010; downloadable in English), the purpose of Finnish pre-primary education is that "the child develops learning-to-learn skills and positive self-image; as well as acquires basic skills, knowledge and capabilities from different areas of learning in accordance with their age and abilities". Understanding of concepts, classification, comparison, and sorting are specified as objectives for early mathematics (p. 11−12). Pre-primary education also aims to develop children's concentration, listening, communication and thinking skills. The children participate in pre-primary educational activities for five days a week, three hours per day. Usually, formal activities include some training for learning letter names and sounds, as well as number symbols. The activities aim also to support social skills: how to follow instructions, how to work in a group, how to co-operate with peers, and how to take care of oneself and one's own responsibilities. In Finland, 97−98 % of

the cohort takes part in the free pre-primary education (The Finnish Ministry of Education and Culture, 2013).

The first pre-test consisted of two tasks for assessing more general and non-domain specific skill levels of the intervention groups (visuo-spatial and phonological working memory); four tasks for assessing early number skills (verbal counting, dot counting fluency, number comparison, basic arithmetic); and one control task unrelated to the intervention (rapid naming). The second pre-test consisted of the aforementioned four number skill tasks and rapid naming. After these tests, the kindergarteners most at-risk (17) were randomly divided into two intervention conditions. Therefore, 9 children (7 boys, 2 girls; mean age = 80.1 months, SD = 4.5) were instructed to practice with GraphoGame Math (GGM group), and 8 children (4 boys, 4 girls; mean age = 78.4 months, SD = 4.1) to practice with Number Race (NR group). At the beginning of the intervention there were no significant differences in the visuo-spatial skills or phonological working memory, early number skills, or the control task between these two groups (see Table 2). Both groups received intensive intervention for 3 weeks, for 10−15 minutes per day. Finally, all children were post-tested using the aforementioned four early number skill tasks and rapid naming. As mentioned earlier, due to the very low accuracy of identifying number symbols, the number comparison task was excluded from our analyses.

Intervention Conditions

Both intervention tools − GraphoGame Math, (in Finnish and Swedish) and Number Race (open source for multiple languages) − are freely available for children, teachers, and parents. An updated version of GGM can be downloaded from an online educational service (www.lukimat.fi), and NR has its own website (http://thenumberrace.com/nr/home.php) from which a detailed user guide can be downloaded.

GraphoGame Math. Originally GraphoGame Math (GGM) has been designed as part of the GraphoGame project at the University of Jyväskylä in Finland (see Richardson & Lyytinen, 2014).

GGM is targeted primarily at children between 6 and 8 years old. The main purpose of the game version we used in the current study was to support acquisition of basic mathematical concepts and skills, such as, dot counting; the correspondence of number word, quantity and number symbol; basic addition; and basic subtraction skills. GGM consisted of several different tasks that were presented in 50 fields of game content, with approximately 1000 items in total. In all trials the child was instructed to respond by choosing the corresponding visual stimulus according to an auditory cue by clicking on the correct item presented on the screen among incorrect alternatives using the mouse's left button.

GGM included tasks in which the exact relations between numbers were practiced. For example, one type of task required the child to identify a correct number neighbour for a verbally presented number word (number before / number after). This activity was intended to strengthen especially the child's verbal number list (see Fuson, 2009; c.f., Wright, 2003), and thus, verbal counting. GGM also aims at practicing object counting and cardinality through tasks in which the child heard a number word (e.g., "four"), and the ball with the corresponding amount of dots (among other balls with different amounts of dots) had to be clicked (here, four dots). Finally, in GGM, basic arithmetic was practiced through tasks in which the child hears a sum (e.g., "five"), and the ball with the corresponding calculation (e.g., 4 + 1) must be clicked. Analogously, basic subtraction was practiced through tasks in which the child hears a difference (e.g., "two"), and the ball with the corresponding calculation (e.g., 4 - 2) must be clicked. Each task in GGM included a time pressure element created by the slow descent of visual objects on the screen. The child needed to choose the corresponding visual stimulus according to an auditory cue before the stimuli (the correct one among the distractors) had fallen down to be eaten by a "pac-man"-like game figure.

The adaptation in GGM was based on gradually increasing complexity of the content (starting from non-symbolic comparisons and continuing to object counting; number concept training; number neighbors activation; symbolic comparisons; and basic arithmetic). Also, the

number range widened gradually (1−3; 1−6; 1−10; 10−20; 20−30), and the better the child

performed the more alternatives (as distractors) appeared on the screen. The adaptation algorithm

aimed at keeping the individual accuracy rate at around 85%, which meant that GGM kept the child

practicing at certain sub-task until the child managed to reach the pre-determined performance

level, before letting the child to proceed for the more demanding training of the next sub-skill.

GGM gave immediate, continuous, and delayed feedback. After a successful trial, the child

heard a sound signalling a correct response. The selected stimulus stopped, and a yellow star outline

appeared, while the incorrect stimuli continued to fall down. After an unsuccessful trial, the child

heard a sound signalling an incorrect response, and the incorrect stimulus stopped, while the correct

stimulus got a green outlining. After a predefined set of trials, the child received feedback according

to the success during the set; this feedback came in the form of butterflies whose colors indicated

the child's accuracy level. The child also saw the total playing time as a progressive bar on the

screen.

Number Race. Number Race (NR) is aimed primarily at children between 5 and 8 years old.

The original purpose of NR was to remediate dyscalculia by enhancing quantity representation

(Wilson, Dehaene, et al., 2006; Wilson, Revkin, et al., 2006). More specifically, Number Race aims

to enhance and automatize number processing, the mental number line, as well as skills in counting,

basic addition, and subtraction (c.f., designers' definitions; www.thenumberrace.com; see "How it

works"). Within the game, the child is instructed to choose the larger of the two quantities presented

visually by concrete objects (coins or coconuts), symbols, or basic addition and/or subtraction

calculations (see also Wilson et al., 2009, p. 227).

The NR has been developed specifically to support the learning of children with MD. In this

study, we sought to examine the specific effects of NR practice on early number skills in

kindergarteners most at-risk for MD. In NR, verbal counting and verbal number lists were implicitly

practiced. After each selection the child made between two presented quantities ("selection

screen"), a race track appeared ("game board screen"). The child needed to click a square on the track whose order on a path corresponded to the number of quantities selected. After that, the game moved the child's character on the track (a non-numerical path) while simultaneously repeating the number words aloud. The child also clicked a square on the track which corresponded to the number of quantities the enemy character received. This was followed by the aforementioned action. Although dot counting was not explicitly practiced, children could use counting or conceptual subitizing in tasks where they were supposed to select the larger of the two presented set of objects (with a range of 1−9). Finally, basic arithmetic was practiced with a similar type of selection task during which the child saw two arithmetic calculations instead of objects/number symbols. The child had to select the one that produced the larger solution. The calculations were presented both as addition (e.g., 2 + 1 vs. 3 + 2); and subtraction (4 - 1 vs. 3 - 2) tasks; or the two task types were mixed (e.g., 2 + 2 vs. 3 - 1).

The adaptation in NR was based on numerical distance, notation, and time pressure being related to the child's performance. For example, the differentiation was supposed to be easier between two distant quantities than between two closer ones (see also Dehaene, 2011, pp. 60−61). As such, in NR, numerical notation changed sensitively between concrete and more complex notations. In terms of the time pressure, after a certain number of successful trials, the enemy character (located on the top of the screen) moved actively for being quicker than the child in reaching the larger amount of the two quantities. In the version of NR we used in our study, the number range in the comparison varied from 1 to 9, and each race track consisted of 40 steps. The item selection algorithm of the game tried to keep the probability of success above 75%.

NR also gave immediate, continuous and delayed feedback. Every time the child managed to choose the bigger quantity, the child heard the sound of applause; conversely, if the child chose the smaller quantity, the child heard a short sound signalling an incorrect response. Every time the

child won a single track, the child could unlock a fish (underwater) or a butterfly (jungle). If the child won many tracks, the child was allowed to unlock new characters to use for playing.

Cognitive Skills Measures Administered in the First Pre-Test

Visuo-spatial Working Memory. The Corsi blocks task is a widely used test designed to assess visuo-spatial working memory (Corsi, 1972; Milner, 1971). A board (8x10 inch) with wooden cubes (1.25 inch) comparable to the original test was used. The child was asked to touch the cubes in the same serial order according to a given model. The span increased by one after every two sets. If the child gave two consecutive incorrect responses, the testing was discontinued. For each set the child correctly repeated, one point was awarded (for a maximum of 16 points). The sum was used in the analyses. Cronbach's alpha for the Corsi blocks tapping task has been found to be .61 (e.g., Busch, Farrel, Lisdahl-Medina, & Krikorian, 2005).

Phonological Working Memory. The Nonword repetition task from the Neuropsychological tests for Children (NEPSY; Korkman, Kirk, & Kemp, 2008) was used to assess phonological working memory. In this task, the child was asked to repeat non-words, which were orally given by the tester, one at a time. There were 16 items that increased in length and complexity. If the child gave four consecutive incorrect responses, the testing was discontinued. The score was the number of correctly repeated items. The sum score was included in the analyses. Cronbach's alpha in this test has been found to be .71 (Korkman, 2000).

Early Number Skills Measures Administered in Two Pre-Tests and Post-Test

Verbal Counting. Verbal counting skills were measured by three separate verbal counting tasks adapted from the Early Numeracy Test (Van Luit, Van de Rijt, & Aunio, 2006). In the counting forward subtest the child was asked to count forward starting from number 1. For correctly reaching the number words 2−9 one point was awarded. For reaching the number words 10−19 two points were awarded. For reaching 20−23 three points, and reaching 24 four points were awarded. In the

counting backward subtest, the child was asked to count backward from 15. The number words for 15, 14 and 13 were given as a model by the tester. If the child was able to count backward correctly until the number words 12−10, the child received one point; reaching 9−6 two points; 5−2, three points; and reaching 1 four points were awarded. In the skip counting subtest, the child was asked to count every second number word starting from 2. The tester provided number words 2, 4, and 6 as a model to begin. If the child was able to continue to the number words 8 or 10 one point was awarded; to 12, 14, 16, or 18 two points; to 20 or 22, three points; and to 24 four points were awarded. A sum score of these three subtests (for a maximum of 12 points) was used in the analyses. Cronbach's alpha was .79 in the first, and .78 in the second pre-test. The Spearman correlation coefficient for test−retest in the sample ($n = 47$) was .73, $p < .001$ (two-tailed).

Object counting. Object counting was assessed by a task in which one to six black randomly arranged dots were presented on a computer screen. The child was asked to say the number of dots aloud as quickly as possible. If the child responded correctly, the tester clicked the mouse's left button. If the child responded incorrectly, the tester clicked the mouse's right button. This test consisted of 4 practice items and 18 test items, with three presentations of one- to six-dot items each. The number of correct responses and reaction times were scored. Because the accuracy of recognizing the dots was over 85% for every task in every assessment point among all participants, the accuracy score was excluded from the analyses. The median of the reaction times for each dot group (1−6) was used for computing two variables. Subitizing fluency (the mean of median reaction times for correctly recognizing dot groups 1−3) was used as a variable according to earlier studies (see Bartelet, Ansari, Vaessen, & Blomert, 2014; subitizing range). Cronbach's alpha was .76 in the first, and .85 in the second pre-test. The other variable used was dot counting fluency (the mean of median reaction times for correctly recognizing dot groups 4−6) based on Bartelet, Ansari, and colleagues (2014; counting range). Cronbach's alpha was .63 in the first, and .60 in the second pre-

test. The Spearman correlation coefficient for test−retest in the sample ($n = 47$) was .76, $p < .001$ (two-tailed) in subitizing fluency, and .66, $p < .001$ (two-tailed) in dot counting fluency.

Basic Arithmetic. Basic arithmetic was measured by a paper and pencil test consisting of two parts: 1) concrete object counting (3 tasks) and 2) symbolic calculation parts (28 tasks). The symbolic calculations included tasks like the following: 2 + 1 = __; 4 − 1 = __; 7 + __ = 14; 15 − __ = 9; 3 + 4 + 6 = __; __ − 3 = 10; 16 = 9 + __ (Aunola & Räsänen, 2007). The test began with the symbolic calculations. The child was instructed to resolve as many of the problems as possible in 3 minutes. A stopwatch was used to measure time. If the child could not solve any calculation items, the child was asked to count objects (circles and squares) from three separate pictures and to add the corresponding number symbol next to each picture. The test was originally developed for the longitudinal data collection and thus, the test included multiple arithmetic combinations for avoiding ceiling effect in later primary school grades. The score for basic arithmetic skills was the sum of correct responses. Those who managed to calculate at least one symbolic problem were automatically given three points for object counting. The maximum score was 31 (3 + 28). The Spearman correlation coefficient for test−retest in the sample ($n = 47$) was .82, $p < .001$ (two-tailed).

Control Measure Administered in Two Pre-Tests and Post-Test

Rapid Naming. The test of Rapid serial naming (RAN) of colors (Denckla & Rudel, 1974; standardized Finnish version by Ahonen, Tuovinen, & Leppäsaari, 2006) was included in all three assessments to control for the specificity of the intervention effects. RAN consisted of five colored squares (black, red, yellow, green, and blue) each repeated several times in pseudorandom order, with no consecutive presentations of the same color. Altogether 50 stimuli were arranged in five rows. Before the test, practice items were presented to ensure that the child knew the names of colors. The child was instructed to name all stimuli as quickly and accurately as possible. A stopwatch was used to measure the time for completion, which was used in the analyses. The

Spearman correlation coefficient for test−retest in the sample ($n = 47$) was .83, $p < .001$ (two-tailed).

Procedure

The children were assessed individually at each of the following three time points: February, February−March, and April. Each assessment session was held in a quiet, separate room in the day care center and lasted approximately 20 to 30 minutes.

After two pre-tests, the children were randomly allocated into two intervention conditions, practicing with either GGM or NR. The children were instructed to play individually with their headphones on (without tutoring) for 12 to 15 times in a 3-week period during their kindergarten hours. Each session was instructed to last 10 to 15 minutes. The study aimed for a minimum exposure to practice time at 120 minutes, which was realized in both intervention conditions. The kindergarten teachers organized the intervention sessions, and helped the children to log in, and log out of the intervention games. To assess intervention fidelity, the teachers also reported the number and length of each session in a practice diary.

Data Analyses

The average scores of two pre-tests of each early number skill (verbal counting, subitizing fluency, dot counting fluency and basic arithmetic) and the control (RAN) measure were used as the initial level score. The Corsi blocks task and the Non-word repetition task were measured once in the first pre-test.

The analysis was made using SPSS version 20. Non-parametric methods were used for the analyses because the variables were not normally distributed and the sample sizes were small (*GGM* = 9, *NR* = 8). Therefore, the Wilcoxon signed-rank test was used to analyze within-group intervention effects. The results were interpreted with exact, one-tailed *p* values. To calculate the within-group effect sizes of the Wilcoxon signed-rank test, the following formula was used: ES (*r*)

= Wilcoxon $Z/\sqrt{N}$, where N is the number of observations (Field, 2013). The between-group differences in the initial level, the intervention gain scores, and the total exposure to intervention were analyzed by the Mann-Whitney U-test. Here the results were interpreted with exact, two-tailed p values. Table 2 presents the intervention group averages at the initial and post-test levels, as well as the significant within-group gain scores.

RESULTS

The effect of intervention for the GGM group on verbal counting was statistically significant, Wilcoxon $Z = -1.95$, $p = .031$, $r = .46$ (Table 2). There was also a significant intervention effect in dot counting fluency, Wilcoxon $Z = -2.19$, $p = .014$, $r = .52$ (Table 2). Altogether 6 children of 9 achieved higher raw scores in verbal counting, and 8 children of 9 were more fluent in dot counting after the intervention. The child with a slower speed in the post-test improved the most in accuracy. Overall, the significant change in dot counting fluency did not result from a lower accuracy level in the post-test. In contrast, the children retained their accuracy level, or were more accurate in the post-test. There was no significant improvement in basic arithmetic, or control task, rapid naming.

For the NR group, a significant intervention effect was seen in basic arithmetic, Wilcoxon $Z = -2.53$, $p = .008$, $r = .63$ (Table 2). Altogether 6 children of 8 achieved higher raw scores in basic arithmetic after NR practice. There was no significant improvement in verbal counting, dot counting fluency, or control task, rapid naming.

Finally, there was a significant between-group difference in intervention gain scores of basic arithmetic ($U = 13.0$, $Z = -2.30$, $p = .014$, $r = .56$), favoring the NR group. There were no other between-group differences in gain scores of early number skills or rapid naming.

The fidelity of intervention was satisfactory in both groups, with all participants reaching the target of 120 practice minutes. The total exposure times ranged from 142 minutes (approx. 9.5 minutes per day) to 237 minutes (approx. 15.8 minutes per day) in the GGM group and from 169

minutes (approx. 11.3 minutes per day) to 350 minutes (approx. 23.3 minutes per day) in the NR group. One child who played 350 NR minutes was an outlier in the sample. After excluding the outlier, the exposure times in the NR group ranged from 169 to 260 minutes (approx. 17.3 minutes per day). However, the outlier was included in non-parametrical analyses based on rank-orders instead of mean scores.

When comparing the exposure to intervention between the GGM and NR groups, there was no significant difference in the number of sessions practiced (GGM: $M = 10.78$, $SD = 2.05$, $MD = 11.00$; NR: $M = 11.25$, $SD = 1.83$, $MD = 11.00$; Mann-Whitney $U = 30.5$, $Z = -0.55$, $p = .619$). However, the difference in total playing time in minutes reached significance ($U = 15.5$, $Z = -1.97$, $p = .049$), indicating longer playing times for the NR group (GGM: $M = 188.00$, $SD = 26.33$, $MD = 190.00$; NR: $M = 232.14$, $SD = 56.31$, $MD = 220.50$). This difference might be due to instructions for playing NR: the children were instructed to end their session only after finalizing an uncompleted race track (from the start point to finish). This instruction was given to ensure that their progress would be recorded per each session. Because of differences in exposure times, and because of the different numerical processes built-in to the two games, the association between playing times and intervention gain scores was analyzed within sub-groups. The results indicated a non-significant correlation (Spearman's rho) between GGM minutes played and gain scores of verbal counting (.29) and between GGM minutes played and gain scores in dot counting fluency (.08). There also was no significant correlation between sessions played and the aforementioned gain scores (.49; .03, respectively). In the NR group, there was no significant correlation between NR minutes played and gain scores in basic arithmetic (.55), or between sessions played and gain scores in basic arithmetic (.10).

Insert Table 2 here.

DISCUSSION

The main purpose of this study was to examine if short and intensive practice with mathematical computer programs can support early number skills in kindergarteners most at-risk for MD. The results indicated a significant intervention effect for verbal counting and dot counting fluency when the children practiced with GraphoGame Math (GGM), and in basic arithmetic when they practiced with Number Race (NR). The effect sizes were relatively large for all improvements at group level ($r = .46$; $.52$; $.63$, respectively; c.f., Cohen, 1992). Between-group difference was found in gain scores of basic arithmetic, favoring the NR group.

It is unlikely that the practice effects found in both intervention groups were due to maturation, kindergarten teaching, or test−retest effect because the intervention period was short, the found intervention effects were group specific, and the test−retest effect was controlled for with two pre-tests. Moreover, if the observed improvements were due to domain general factors, parallel gains could be expected in the control measure since all assessment tasks shared the fluency requirement. However, there were no within-group effects in the control measure (rapid naming).

The improvements in verbal counting and dot counting fluency in the GGM group can be explained with the nature of the GGM practice itself: it focused on exact discriminations of numerical representations. Also the time limit in each trial encouraged fluency. In addition to counting, GGM practice might have strengthened the concept of cardinality, the relations between numbers (number neighbors, number comparison), and the ability to detect quickly the sub-groups of objects. For example, in each dot counting trial, the child needed to count different quantities in order to pick out the correct stimulus among incorrect ones within a limited time. Therefore, GGM could have directed the children towards using faster and more efficient strategies in determining the number of objects. This would mean seeing a set of five dots as a combination of three-and-two instead of counting the dots one by one. This would reflect conceptual subitizing (Sarama & Clements, 2009).

The results are encouraging since verbal counting is shown to have a strong connection to later arithmetic at school age (e.g., Desoete & Grégoire, 2006; Koponen et al., 2013; Lepola et al., 2005). There is also evidence that fluency in object counting is rather stable between different fluency-level groups (e.g., Reeve et al., 2012). In previous CAI studies, positive effects on object counting accuracy in low-performing children have been reported (Ortega-Tudela & Gómez-Ariza, 2006; Table 1). At-risk children also have benefited from computerized practice (Elliot & Hall, 1997; Table 1). The latter gain was seen on a larger achievement test (TEMA-2) containing object counting as one sub-skill. Hence, it is difficult to conclude whether or not the gain in Elliot's and Hall's study (1997) resulted specifically from an improvement in object counting, or if it simply reflected general improvements in all sub-skills.

Also, a significant effect of intervention was found in basic arithmetic in the NR group. In earlier NR studies, positive effects on arithmetic skills have been found in school-aged children with specific MD status (Wilson, Revkin, et al., 2006, the original version of NR) and without it (see Obersteiner, Reiss, & Ufer, 2013; two different experimental versions of NR were used). The effect is logical considering the content of the game. In NR, the quantities are first presented as concrete objects and number symbols, but quite soon also as basic addition and/or subtraction calculations (ongoing adaptation in notation, numerical distance, and time pressure). As an example, after the child has chosen a calculation like 3 + 2, NR repeats aloud "you chose - three plus two - equals five" while simultaneously presenting all symbols "3 + 2 = 5". This might help children to learn the association between verbal (spoken) and written (numbers and arithmetical symbols i.e. plus, minus, equal) representations of basic addition and/or subtraction calculations. This finding is encouraging since early number combination and story problem solving skills seem to predict later calculation procedures and applied problem solving (e.g., Jordan et al., 2009), and the effect size was relatively large. In earlier CAI studies (other than NR), positive intervention

effects in basic addition have been found in groups of at-risk children with performance below the 25th percentile (e.g., Baroody et al., 2012, 2013; Fuchs et al., 2006).

Furthermore, the between-group comparison revealed a significant difference in gain scores of basic arithmetic, favoring the NR group. This result could be due to structure of the games. In NR, basic addition and/or subtraction calculations are presented quite soon after concrete objects and number symbols since the numerical notation is adaptive for accuracy. Therefore, the content varies more continuously in NR than in GGM, which is divided into different levels. This type of variation might mean that the children are exposed to arithmetic practice regardless of the number of NR minutes or sessions played. Such an ongoing sensitivity in adaptation of NR could explain the between-group difference in gain scores on basic arithmetic. Indeed, NR focuses on the approximate numerical processes for determining which of the two arithmetic calculations should be selected for receiving a larger amount of objects; however, the children might have needed to estimate in more detail (or even calculate) the sums and differences of the presented calculations. In GGM, by contrast, the numerical content was organized so that basic arithmetic was hierarchically the highest sub-skill practiced; thus, it is possible that the children did not reach the highest training level during the intervention period. GGM had an adaptation that kept the child practicing specific sub-skills until the satisfied performance level was achieved. Additionally, arithmetic practices were perhaps too complex in the current GGM version, even if it did expose the certain basic concepts. In GGM, the child heard a sum, and the correct calculation must be selected among a number of alternatives. In GGM's arithmetic tasks, the operation symbols (plus, minus) were visually presented, but the symbols were not verbally presented at all, unlike in NR. Therefore, both treatments focused merely on procedurally oriented addition and subtraction training. It is probable that such practice should come only after the conceptually oriented training in this target group. As suggested, a good conceptual knowledge allows an efficient application of calculation procedures (Dowker, 2009).

We also examined if the intervention benefit was related to the intervention exposure. The results revealed that the gain scores of verbal counting and dot counting fluency did not correlate significantly with the amount of GGM sessions or minutes played. Despite the significant improvement in basic arithmetic in the NR group, the gain score was not related to total NR sessions or minutes played either. This finding could most likely be explained by the adaptation, which individualizes the practice in both games. The success rate is approximately 85 % in GGM (after achieving a certain amount of correct answers, the child is allowed to move to the next game level); and in NR, the content varies frequently depending on the child's performance. For this reason, there may be variation in children's exposure times for different sub-skill training. Perhaps this variation explains why significant correlations were not found between minutes played, sessions practiced, and certain intervention gains.

In sum, the results of this study are in line with some earlier studies in which CAI has been shown to be effective especially for children with weak skills (e.g., Li & Ma, 2010) over short, intensive practice periods (e.g., Kroesbergen & Van Luit, 2003; Kulik & Kulik, 1991). There are also suggestions that a well-planned adaptive practice is able to identify children's strengths and weaknesses as well as fill their individual gaps (Fuchs, 2005; Fuchs, Fuchs, & Compton, 2012), also when offered in computerized format (Hasselbring, 1986; Slavin & Lake, 2008). Hence, it seems reasonable to offer specific number skill training for kindergarteners most at-risk for MD, especially because it is known that early difficulties tend to be very persistent within this group (e.g., Geary et al., 2008; Morgan et al., 2009; Murphy et al., 2007). Furthermore, it might be worth noting that individually targeted practice with computers allows teachers to concentrate on other methods to enhance children's learning (Clements, 2002), and to enrich their experience.

The specific features of the used games might explain different types of effects observed. GGM focused on exact numerical processes and cardinality, and had an effect on verbal counting and dot counting fluency. On the other hand, NR focused on approximate numerical processes,

which means that neither the cardinality nor exact dot counting was practiced explicitly. After each quantity selection, the game moved characters along a track while the number words were simultaneously repeated which supported directly verbal counting (counting forward) (see the images in Wilson et al., 2009, p. 227). However, the number range was perhaps too concise (1−9) in NR to produce effects for the used verbal counting measures (range 1−24). As already discussed, both assessed and trained arithmetic tasks having a procedural orientation (instead of a conceptual one) might have had an effect on the effect sizes observed in a group of most at-risk for MD.

As mentioned, the adaptation in both games increased the variation in practiced content. Children might have been exposed somewhat differently to specific sub-skill training within a short intervention period of three weeks. It is also possible that our assessment tools were not sensitive enough to pick up development for all skills assessed. As noted earlier, the number comparison task used emphasized speeded responding and was unfit for the most at-risk participants of the current study. There was also a floor effect in basic arithmetic task due to its original purpose in a longitudinal data collection for avoiding a ceiling effect in later primary school grades. It is obviously not straightforward to create assessment materials responsive enough for the whole range of early number skill levels. On the other hand, standardized assessment tools usually are not specific enough for the targeted training contents. As such, this study should be considered a preliminary approach for assessing the intervention effects in children most at-risk for MD.

There are also other limitations in the study. The small sample sizes, like we had in this study, might create a lack of power for revealing less robust effects of the interventions. Our inclusion criterion was stricter than typically used: the children with a performance below the 10th percentiles in verbal counting were included. Obviously, this limited scope means that the results should be interpreted with caution and await replications with larger samples. In further studies, an experimental design with a business-as-usual control group would also be useful in determining the specific effects of the intervention.

In a number of studies it has been pointed out that within small samples the effect sizes tend to be relatively large (e.g., Cheung & Slavin, 2013; Slavin & Lake, 2008). For this reason, instead of effect size values, different benchmarks for interpreting the effectiveness (relevant to intervention, target population, and outcome measures) should be used (Grissom & Kim, 2012; Hill, Bloom, Black, & Lipsey, 2008). In our case, the effect size comparisons were difficult to carry out since the inclusion criterion was stricter, and intervention duration and intensity differed in the current study as compared to earlier CAI studies (see Table 1). Although the effect sizes exceeded the long-time averages presented in meta-analyses (see Räsänen, in press), they should be evaluated with caution. In addition, standardized tests have been suggested for identifying target children (Mazzocco, 2005) and for measuring outcomes (Slavin & Lake, 2008) when studying the real transfer benefit of interventions, or for proposing any method as an evidence-based practice (e.g., Cook, Tankersley, & Landrum, 2009). Nonetheless, in this study, the practice targeted the specific skills of a particular group of children in need of early support. In other words, the purpose of this study was not to evaluate the transfer effects or to compare the effects to the normally-developing reference group that already performs at the ceiling in many prerequisite early number tasks. Finally, it would be useful to have an access to game log data, and to conduct long-term post-assessments.

Some practical implications in terms of developing educational games are worth discussing. Even though it seems that a short and intensive computerized practice can produce condition specific effects with regard to the specific group of children most at-risk for MD, developers should carefully focus on coherent intervention principles. The practice should include explicit instructions; step-by-step procedures; simultaneous training for both concepts and concrete operations; immediate, continuous, and delayed feedback; a motivating environment; and ongoing assessment (cf. Baker, Gersten & Lee, 2002; Fuchs et al., 2008; Gersten et al., 2009). Both intervention programs used in the current study (GraphoGame Math and Number Race) cover the

majority of the aforementioned principles, but there is room for improvement. This means not only a sufficient and rich numerical content with multidimensional task types, but also a carefully planned MD-appropriate user interface; content-based adaptations; and a more pedagogical feedback system. Due to a persistent nature of deficits in arithmetic the early intervention should strengthen both incipient number skills and basic addition/subtraction in a meaningful way. The conceptual basis should be practiced before more procedurally oriented training starts. All of these aforementioned components could have specific influences on the desired immediate and long-term effects. To develop such an appropriate and sensitive tool, multidisciplinary efforts are needed, including mathematics education and psychology researchers, game developers, and big log-data statisticians.

As a recommendation for further studies, the effectiveness of CAI in children most at-risk for MD should be examined with larger samples. As studies have shown individual variation in intervention responsiveness to be large in a group of children with MD (e.g., Dowker & Sigley, 2010; Fuchs et al., 2012; Geary, 2011), we would recommend examining the effects of a tailored, targeted training based on qualified screening assessments. The potential intervention benefits also should be followed by delayed assessments (e.g., Fuchs et al., 2006; Wilson et al. 2009). In addition, game log data could provide more detailed information on an individual level, as well as generate insight into how children actually act while using the programs (see Käser et al., 2011 for more on using log file data in analyses of CAI effectiveness). This type of data could help determine the individual patterns of development within the CAI and give a deeper understanding of the vital factors in producing better learning.

References

Ahonen, T., Tuovinen, S., & Leppäsaari, T (2006). *Nopean sarjallisen nimeämisen testi* (2nd ed.). Jyväskylä, Finland: Niilo Mäki Institute.

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20*(5), 427−435. doi: 10.1016/j.lindif.2010.06.003

Aunola, K., Leskinen, E., Lerkkanen, M-K., & Nurmi, J-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*(4), 699–713. doi: 10.1037/0022-0663.96.4.699

Aunola, K., & Räsänen, P. (2007). *The 3-minutes Basic Arithmetic Test.* Niilo Mäki Institute. Unpublished.

Baker, S., Gersten, R., & Lee, D-S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal, 103*(1), 51−73. Retrieved from http://www.jstor.org/stable/1002308

Baroody, A. J. (1987). The development of counting strategies for single-digit addition. *Journal for Research in Mathematics Education, 18*(2), 141−157. http://www.jstor.org/stable/749248

Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2012). Fostering at-risk kindergarten children's number sense. *Cognition and Instruction, 30*(4), 437–470. doi: 10.1080/07370008.2012.720152

Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2013). Can computer-assisted discovery learning foster first graders' fluency with the most basic addition combinations? *American Educational Research Journal, 1,* 1−41. doi: 10.3102/0002831212473349

Bartelet, D., Ansari, D., Vaessen, A., & Blomert, L. (2014). Cognitive subtypes of mathematics learning difficulties in primary education. *Research in Developmental Disabilities 35*(3), 657−670. doi: 10.1016/j.ridd.2013.12.010

Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D., (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology 117*, 12–28. doi: 10.1016/j.jecp.2013.08.010

Becker, H. J. (1992). Computer-based integrated learning systems in the elementary and middle grades: A critical review and synthesis of evaluation reports. *Journal of Educational Computing Research, 8*(1), 1−41. doi: 10.2190/23BC-ME1W-V37U-5TMJ

Busch, R. M., Farrell, K., Lisdahl-Medina, K., & Krikorian, R. (2005). Corsi block-tapping task performance as a function of path configuration. *Journal of Clinical and Experimental Neuropsychology, 27*(1), 127−134. doi: 10.1080/138033990513681

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*, 1049−1053. doi: 10.1126/science.1201536

Chambers, J. A., & Sprecher, J. W. (1980). Computer Assisted Instruction: Current Trends and Critical Issues. *Communications of the Association for Computing Machinery, 23*(6), 332−342. doi: 10.1145/358876.358881

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88−113. doi: 10.1016/j.edurev.2013.01.001

Chong, S. L., & Siegel, L. S. (2008). Stability of computational deficits in math learning disability from second through fifth grades. *Developmental Neuropsychology, 33*(3), 300–317. doi: 10.1080/87565640801982387

Christmann, E. P., & Badgett, J. L. (2003). A meta-analytic comparison of the effects of computer-assisted instruction on elementary students' academic achievement. *Information Technology in Childhood Education Annual, 1*, 91–104. Retrieved from http://www.editlib.org/p/18866

Clements, D. H. (2002). Computers in early childhood mathematics. *Contemporary Issues in Early Childhood, 3*(2), 160−180. doi: org/10.2304/ciec.2002.3.2.2

Clements, D. H. (2004). Major themes and recommendations. In D. H. Clements, & J. Sarama (Eds.), *Engaging young children in mathematics, standards for early childhood mathematics education* (pp. 7–72). Mahwah, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 122*(1), 155−159. doi: 10.1037/0033-2909.112.1.155

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children, 75*(3), 365–383. Retrieved from http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=3&sid=75a11115-bab9-4bfd-b9f4-18b70eb30267%40sessionmgr115&hid=118

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34*(2).

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* Oxford, UK: University Press.

Dencla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colours, letters and numbers by normal children. *Cortex, 10*(2), 186−202.

De Smedt, B., Noël, M-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*, 48–55. doi: 10.1016/j.tine.2013.06.001

Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences, 16*, 351–367. doi: 10.1016/j.lindif.2006.12.006

Din, F. S., & Calao, J. (2001). The effects of playing educational video games on kindergarten achievement. *Child Study Journal, 31*(2), 95–102. http://clem.mscd.edu/~sandersc/3310%20ed%20video%20article.pdf

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development, 24*(4), 401−410. doi: 10.1016/j.cogdev.2009.09.005

Dowker, A., & Sigley, G. (2010). Targeted interventions for children with arithmetical difficulties. *The British Psychological Society, Monograph Series 2*(7), 65–81. doi:10.1348/97818543370009X12583699332492

Elliot, A., & Hall, N. (1997). The impact of self-regulatory teaching strategies on "at-risk" preschoolers' mathematical learning in a computer-mediated environment. *Journal of Computing in Childhood Education, 8*(2/3), 187–198.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: SAGE Publications Ltd.

Finnish Ministry of Education and Culture. (2013). Selvitys esiopetuksen velvoittavuudesta [Report of obligation to participate in Finnish kindergarten curriculum], Finland: Ministry of Education and Culture.

Finnish National Board of Education. (2010). Core curriculum for preschool education in Finland. Downloadable (in English) from http://www.oph.fi/download/153504_national_core_curriculum_for_pre-primary_education_2010.pdf

Fuchs, L. S. (2005). Prevention research in mathematics: Improving outcomes, building identification models, and understanding disability. *Journal of Learning Disabilities, 38*(4), 350–352. doi: 10.1177/00222194050380041201

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2012). The early prevention of mathematics difficulty: Its power and limitations. *Journal of Learning Disabilities, 45*(3), 257–269. doi: 10.1177/0022219412442167

Fuchs, L. S., Fuchs, D., Hamlet, C. L., Powell, S. R., Capizzi, A. M., & Seethaler, P. M. (2006). The effects of computer-assisted instruction on number combination skill in at-risk first graders. *Journal of Learning Disabilities, 39*(5), 467–475. doi: 10.1177/00222194060390050701

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008).

  Intensive intervention for students with mathematics disabilities: Seven principles of

  effective practice. *Learning Disability Quarterly, 31*(2), 79–92. doi: 10.2307/20528819

Fuson, K. C. (1988). *Children's counting and concepts on number.* New York: Springer-Verlag.

Fuson, K. C. (2009). Avoiding misinterpretations of Piaget and Vygotsky: Mathematical teaching

  without learning, learning without teaching, or helpful learning-path teaching? *Cognitive*

  Development, 24, 343–361. doi: 10.1016/j.cogdev.2009.09.009

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1),

  4–15. doi: 10.1177/00222194040370010201

Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical

  learning disabilities and persistent low achievement in mathematics. *Journal of*

  Developmental and Behavioral Pediatrics, 32(3), 250–263.

  doi:10.1097/DBP.0b013e318209edef

Geary, D. C. (2013). Learning disabilities in mathematics: Recent advances. In H. L. Swanson, K.

  Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (second edition) (pp.

  239−255). New York: Guilford Press.

Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children

  and children with mathematical learning disability. *Journal of Learning Disabilities, 45*(4),

  291−307. doi: 10.1177/0022219410392046

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line

  representations in children with mathematical learning disability. *Developmental*

  Neuropsychology, 33(3), 277−299. doi: 10.1080/87565640801982361

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA:

  Harvard University Press.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. R. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202–1242. doi: 10.3102/0034654309334431

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

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). New York, NY: Routledge.

Hasselbring, T. S. (1986). Research on the effectiveness of computer-based instruction: a review. *International Review of Education, 32*(3), 313–324. doi: 10.1007/BF02426065

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177. doi: 10.1111/j.1750-8606.2008.00061.x

Howard, J. R., Watson, J. A., Brinkley, V. M., & Ingels-Young, G. (1994). Comprehension monitoring, stylistic differences, pre-math knowledge, and transfer: A comprehensive pre-math/spatial development computer-assisted instruction (CAI) and LOGO curriculum designed to test their effects. *Journal of Educational Computing Research, 11*(2), 91–105. doi: 10.2190/8WK8-Q6T1-2V0R-K2RW

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice, 22*(1), 36–46. doi: 10.1111/j.1540-5826.2007.00229.x

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

Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J-E. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skills. *Journal of Experimental Child Psychology, 97*, 220–241. doi: 10.1016/j.jecp.2007.03.001

Koponen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology, 105*(1), 162–175. doi: 10.1037/a0029285

Korkman, M. (2000). *Nepsy. Käsikirja II.* Helsinki, Finland: Psykologien Kustannus Oy.

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY. A developmental neuropsychological assessment manual*, Finnish Edition: NEPSY II - Lasten neuropsykologinen tutkimus (Korkman, Kirk, & Kemp, 2008). Helsinki, Finland: Psykologien Kustannus Oy

Kroesbergen, E. H., & Van Luit, J. (2003). Mathematics interventions for children with special educational needs. A meta-analysis. *Remedial and Special Education, 24*(2), 97–114. doi: 10.1177/07419325030240020501

Kulik, C-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*, 75–94. doi: 10.1016/0747-5632(91)90030-5

Käser, T., Kucian, K., Ringwald, M., Baschera, G. M., von Aster, M., & Gross, M. (2011). Therapy software for enhancing numerical cognition. In J. Özyurt, A. Anschütz, S. Bernholt, & J. Lenk (Eds.), *Interdisciplinary perspectives on cognition, education and the brain* (pp. 207−216). Oldenburg, BIS-Verlag. (Hanse-Studies Vol. 7)

Lepola, J., Niemi, P., Kuikka, M., & Hannula, M. M. (2005). Cognitive-linguistic skills and motivation as longitudinal predictors of reading and arithmetic achievement: A follow-up study from kindergarten to grade 2. *International Journal of Educational Research, 43*, 250–271. doi: 10.1016/j.ijer.2006.06.005

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students'

    mathematics learning. *Educational Psychology Review, 22*(3), 215−243. doi:

    10.1007/s10648-010-9125-8

Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation

    fluency in second grade. *Journal of Learning Disabilities, 41*(5), 451–459. doi:

    10.1177/0022219408321126

Mazzocco, M. M. M. (2005). Challenges in identifying target skills for math disability screening

    and intervention. *Journal of Learning Disabilities, 38*(4), 318−323. doi:

    10.1177/00222194050380040701

Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning

    disability. *Learning Disabilities Research & Practice, 20*(3), 142–155. doi: 10.1111/j.1540-

    5826.2005.00129.x

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-

    related skills on academic trajectories at the end of elementary school. *Early Childhood*

    Research Quarterly, 21, 471–490. doi: 10.1016/j.ecresq.2006.09.003

McCollister, T. S., Burts, D. C., Wright, V. L., & Hildreth, G. J. (1986). Effects of computer-

    assisted instruction and teacher-assisted instruction on arithmetic task achievement scores of

    kindergarten children. *Journal of Educational Research, 80*(2), 121–125.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in

    man. *British Medical Bulletin, 27*(3), 272−277.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children

    with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*(4), 306−321.

    doi: 10.1177/0022219408331037

Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive

    characteristics of children with mathematics learning disability (MLD) vary as a function of

the cutoff criterion used to define MLD. *Journal of Learning Disabilities, 40*(5), 458−478.

doi: 10.1177/00222194070400050901

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental

representations of number can enhance first-grade students' basic number processing and

arithmetic skills. *Learning and Instruction, 23*, 125−135. doi:

10.1016/j.learninstruc.2012.08.004

Ortega-Tudela, J. M., & Gómez-Ariza, C. J. (2006). Computer-assisted teaching and mathematical

learning in Down syndrome children. *Journal of Computer Assisted Learning, 22*, 298−307.

doi: 10.1111/j.1365-2729.2006.00179.x

Piazza, M., Mechelli,A., Butterworth, B., & Price. C. J. (2002). Are subitizing and counting

implemented as separate or functionally overlapping processes? *NeuroImage, 15*, 435−446.

doi:10.1006/nimg.2001.0980

Polet, J., & Koponen, T. (2011). LukiMat - Oppimisen arviointi: Matematiikan tuen tarpeen

tunnistamisen välineet esikouluun. Tekninen opas. Downloadable from

http://www.lukimat.fi

Praet, M., & Desoete, A. (2014). Enhancing young children's arithmetic skills through non-

intensive, computerized kindergarten interventions: A randomized controlled study.

Teaching and Teacher Education 39, 56−65. doi: 10.1016/j.tate.2013.12.003

Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in

markers of core numerical competencies. *Journal of Experimental Psychology, 4*(141), 649–

666. doi: 10.1037/a0027520

Richardson, U., & Lyytinen, H. (2014). The GraphoGame method: The theoretical and

methodological background of the technology-enhanced learning environment for learning

to read. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*

10(1), 39–60. Downloadable from

http://www.humantechnology.jyu.fi/articles/volume10/2014/Richardson_Lyytinen.pdf

Räsänen, P. (in press). Computer-assisted interventions on basic number skills. In R. Cohen

Kadosh, & A. Dowker (Eds.), *Oxford handbook of numerical cognition*. London: Oxford

University Press.

Räsänen, P., Salminen, J. Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted

intervention for children with low numeracy skills. *Cognitive Development, 24*, 450−472.

doi: 10.1016/j.cogdev.2009.09.003

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning

trajectories for young children.* New York, NY: Routledge.

Seo, Y-J., & Bryant, D. P. (2009). Analysis of studies of the effects of computer-assisted instruction

on the mathematics performance of students with learning disabilities. *Computers &

Education, 53*, 913–928. doi:10.1016/j.compedu.2009.05.002

Skagerlund, K., & Träff, U. (2014). Number processing and heterogeneity of developmental

dyscalculia: Subtypes with different cognitive profiles and deficits. *Journal of Learning

Disabilities.* doi: 10.1177/0022219414522707

Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best evidence

synthesis. *Review of Educational Research, 78*(3), 427−455. doi:

10.3102/0034654308317473

Starkey, P., & Cooper, R. G. Jr, (1980). Perception of numbers by human infants. *Science,

210*(4473), 1033−1035. doi: 10.1126/science.7434014

Van Luit, J. E. H., Van de Rijt, B. A. M., & Aunio, P. (2003). Early Numeracy test, Finnish Edition:

Lukukäsitetesti (Van Luit, Van de Rijt, & Aunio, 2006). Helsinki, Finland: Psykologien

Kustannus Oy.

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game

intervention on accessing number sense in low-socioeconomic-status kindergarten children.

Mind, Brain, and Education, 3(4), 224–234. doi: 10.1111/j.1751-228X.2009.01075.x

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles

underlying the design of "The Number Race", an adaptive computer game for remediation

of dyscalculia. *Behavioral and Brain Functions, 2*(19), 1–14. doi: 10.1186/1744-9081-2-19

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment

of The Number Race, an adaptive computer game for remediation of dyscalculia. *Behavioral

and Brain Functions, 2*(20), 1–36. doi: 10.1186/1744-9081-2-20

Wright, R. J. (2003). Mathematics recovery: A program of intervention in early number learning.

Australian Journal of Learning Disabilities, 8(4), 6–11. doi: 10.1080/1940415030954674

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155−193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive

Psychology, 24*, 220−251.

Table 1

Descriptions of Computer-Assisted Intervention Studies in At-Risk Children

Study	Sample	Age	Status	Time	Sessions	Main result (ES / ES_m)[k]
Baroody et al. (2012)	15+13	5.58	at-risk MD[b]	19	20[h] + 20 x 30 min	Addition (ES_m = .30)
Baroody et al. (2013)	43+21	6.5	at-risk MD[b]	20	20[h] + 20 x 30 min	Addition (n+1/1+n) (ES = .39)
Din et al. (2001)[a]	24+23	5.5	low SES	11	55 x 40 + 30 min[i]	No significant training effect
Elliot et al. (1997)[a]	18+18	4	at-risk LD[c]	6	15 x 20 min	TEMA-2 (ES = .54)
Fuchs et al. (2006)	16+17	6−7	at-risk MD[d]+RD[e]	18	50 x 10 min	Addition (ES = .44)
Howard et al. (1994)	32	4.42	middle SES	9	27 x 15 min	TEMA (ES[l])
McCollister et al. (1986)[a]	28+25	5−6	low / middle SES	< 2	10 x 10−15 min	Numeral recognition (ES = .43)
Ortega-Tudela et al. (2006)	10+8	6.55	low performing[f]	21	15 x 35 min	Enumeration (ES_m = .61)
Praet et al. (2014)	44+39+49	5.67	middle SES	5	8 x 25 min	Arithmetic (ES_m = .39)[m]
Räsänen et al. (2009)	29+15+15	6.55	low performing[g]	3	15 x10−15 min	Symbolic comparison (ES = .22)[n]
Wilson et al. (2009)	27+26	5.6	low SES	14	6 + 4 x 20 min[j]	Symbolic comparison (ES_m = .39)[o]

Note. Sample = Sample sizes (treatment + control) used for analyses. Age = Mean age in years (*SD*s were not always reported). MD = mathematics difficulties. SES = socioeconomic status. LD = learning difficulties. RD = reading difficulties. Time = Time used for practice in weeks. Main result = Main significant training effects. $ES_{(m)}$ = Effect size$_{(mean)}$. TEMA = Test of Early Mathematics Ability (Ginsburg & Baroody, 1983, 1990, 2003).
[a]Training in pairs. [b]In the bottom 25th percentile in TEMA-3 (and/or personal and/or familial risk factors). [c]Identified as being at-risk of early learning difficulties. [d]In the bottom 25th percentile in arithmetic and [e]in the bottom 17th percentile in reading. [f]Not acquired counting and cardinality concepts according to a curriculum-based evaluation. [g]Identified by kindergarten teachers as having a need for numerical support. [h]20 sessions with manual games, 20 with CAI. [i]40 min at kindergarten, 30 min at home. [j]Six sessions with math, four sessions with reading software. [k]If not reported, effect sizes (Cohen's *d*) were calculated from reported statistics by recommended formulas (Cohen, 1992), and thereafter all *d* values were converted to *r* values for being comparable to the current study results. [l]Could not be calculated. [m]Sample sizes of low performers (<25th percentile) in three conditions were not reported. Between-groups *ES*s were calculated for the main sample. [n]Counted as an *ES* for both intervention conditions. [o]Counted from quadratic analyses for the lowest performers.

Table 2

Group Means, Standard Deviations, Medians, and Significant Within-Group Gain Scores, as well as Between-Group Comparisons

		Intervention condition					
		GGM (*n* = 9)		NR (*n* = 8)		Group comparisons
Variable	Time point	*M (SD)*	*MD*	*M (SD)*	*MD*	*Mann-Whitney U*
Cognitive skills						
Corsi blocks (max. 16)	T1	4.67 (1.00)	5.00	5.25 (1.04)	5.00	
Non-words (max. 16)	T1	7.56 (2.70)	9.00	8.25 (1.49)	8.00	
Early number skills						
Verbal counting (max. 12)	T1	**4.17 (1.94)**	4.00	5.81 (1.62)	6.25	
	T2	**6.11^b (3.48)***	7.00	7.50 (3.51)	7.50	
Subitizing fluency (seconds)	T1	1.42 (0.18)	1.38	1.56 (0.27)	1.63	
	T2	1.31 (0.39)	1.23	1.43 (0.36)	1.45	
Dot counting fluency (seconds)	T1	**3.82 (0.44)**	**3.73**	3.88 (0.96)	3.59	
	T2	**3.11^c (0.61)***	**3.24**	3.62 (0.82)	3.58	
	gain	-0.71 (0.71)	-0.96	-0.26 (0.75)	-0.37	
Basic arithmetic (max. 31)^a	T1	3.67 (1.32)	4.00	**4.94 (1.32)**	**4.25**	
	T2	3.89 (1.36)	4.00	**5.75^d (1.17)***	**5.00**	NR > GGM**
	gain	0.22 (0.97)	0.00	0.81 (0.37)	1.00	NR > GGM*
Control task						
Rapid naming (seconds)	T1	72.39 (12.11)	69.50	67.63 (17.44)	65.50	
	T2	64.89 (14.13)	58.00	65.63 (12.98)	65.50	

Note. GGM = Graphogame math group; NR = Number race group; T1 = initial level; T2 = post-test.
Within-group effects of intervention are shown in boldface.
^{a}Time limited task. ^{b}Wilcoxon *ES* (*r*) = .46; ^{c}.52; ^{d}.63.
* *p* < .05. ** *p* < .01.