**Author(s):** Salminen, Jonna; Koponen, Tuire K.; Leskinen, Markku; Poikkeus, Anna-Maija; Aro, Mikko

**Title:** Individual variance in responsiveness to early computerized mathematics intervention

**Year:** 2015

**Version:**

**Please cite the original version:**

Salminen, J., Koponen, T. K., Leskinen, M., Poikkeus, A.-M., & Aro, M. (2015). Individual variance in responsiveness to early computerized mathematics intervention. Learning and Individual Differences, 43, 124-131. https://doi.org/10.1016/j.lindif.2015.09.002

# Individual variance in responsiveness to early computerized mathematics intervention

Jonna B. Salminen, M.Ed., University of Jyvaskyla, Department of Education, Special Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland, jonna.salminen@jyu.fi;

Tuire K. Koponen, PhD, Niilo Mäki Institute, Jyväskylä, PO box 35 40014 Jyväskylä, Finland, tuire.koponen@nmi.fi;

Markku Leskinen, PhD, University of Jyvaskyla, Department of Education, Special Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland, markku.leskinen@jyu.fi;

Anna-Maija Poikkeus, PhD, University of Jyvaskyla, Department of Teacher Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland, anna-maija.poikkeus@jyu.fi;

and

Mikko T. Aro, PhD, University of Jyvaskyla, Department of Education, Special Education, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland, mikko.aro@jyu.fi

Correspondence concerning this article should be addressed to

Jonna Salminen, University of Jyvaskyla, Department of Education, Special Education,

phone number: +358 40 805 3027

e-mail address: jonna.salminen@jyu.fi

postal address: P.O. Box 35, FI-40014 University of Jyvaskyla, Finland

**Individual variance in responsiveness to early computerized mathematics intervention**

We examined the effects of short, intensive computerized intervention in early number skills for kindergarteners with poor addition skills (below 1.5 *SD*). The mathematical content of the software was hierarchically organized, starting from one-to-one correspondence, comparing and ordering, and proceeding via number concept and counting to basic addition. The results showed positive within-group effects for basic addition (Wilcoxon *ES (r)* = .59), verbal counting (.56), and the Number Sets Test (.45; see Geary et al., 2009). The effects remained stable over a 9-week follow-up period. However, there was no significant between-group difference in terms of gain scores as compared to a wait-list control group. Based on game-log data, individual variance in responsiveness to the intervention was analyzed. Even though the findings suggest that adaptive, hierarchically organized content could provide effective support for some children with poor early number skills, more specific instruction and feedback system are needed in individualizing interventions.

*Keywords:* computer-assisted intervention (CAI); kindergarten; early number skills; game-log data; response to intervention (RTI)

## 1. Introduction

Basic arithmetic (i.e., addition and subtraction skills) is an important predictor for later school mathematics achievement above and beyond the influence of intelligence (Duncan et al., 2007; Geary, 2011a; Geary, Hoard, & Bailey, 2012). Difficulties in basic arithmetic and fact retrieval are very persistent (Geary, 2011b), and they constitute a core feature of mathematics difficulties (MD) (Gersten, Jordan, & Flojo, 2005). Accumulated knowledge about the development of early number skills as a basis for arithmetic has helped to advance means of early identification and support (Butterworth, 2005; Morgan, Farkas, & Wu, 2009). However, there is still a need to examine effective, evidence-based intervention methods (Butterworth, Varma, & Laurillard, 2011; Jordan & Levine, 2009), especially among kindergarteners with poor early number skills (performance below 10th percentile; Morgan, et al., 2009; Murphy, Mazzocco, Hanich, & Early, 2007).

Early arithmetic skill seems to develop via different hierarchical modules of sub-skills. In this regard, approximate magnitude discrimination is the innate skill of subitizing small sets of objects and quickly differentiating which of the two sets of objects is the larger when the difference between the quantities is significant enough (Geary, 2013; Dehaene, 2011). The ability to recite number words evolves later with the development of expressive language skills, which thus enables meaningful counting (Krajewski & Schneider, 2009). After that, the link between the number words, quantities and symbols becomes precise (Geary, 2013; Krajewski & Schneider, 2009). This developmental step is required for describing the exact amount of quantities exceeding the subitizing range. Furthermore, an explicit number system, including an understanding of the relationships between numbers, is the next vital step for composing and decomposing, and thus, basic arithmetic skills (cf. Geary, 2013; Krajewski & Schneider, 2009). Von Aster (2000) has noted that individual development is also dependent on the maturation of semantic, verbal, and visual/symbolic modules, which are all necessary for calculation skills.

Due to the diversity of the deficits associated with MD (e.g., Rubinsten & Henik, 2009), as well as the heterogeneity among individuals with MD (e.g., Geary, 2004; Jordan, Hanich, & Kaplan, 2003; Von Aster & Shalev, 2007), there is a call for tailored interventions (Dowker, 2001; Geary, 2011b; Slavin & Lake, 2008), and continuous evaluation and identification of children who are not responding to support (Fuchs, Fuchs, & Compton, 2012). Therefore, adaptive intervention programs with dynamic, simultaneous assessment tools could be beneficial in identifying children with (or at-risk for) MD, and for assisting teachers in planning individualized support.

A well-planned computer-assisted intervention (CAI) offers several possibilities for tailored practice (Seo & Bryant, 2009; Slavin & Lake, 2008), even though the main trend in terms of the effectiveness of CAI on number skills has been suggested to be in decline in recent decades (Cheung & Slavin, 2013). The factors behind CAI’s positive effects on arithmetic development are challenging to identify due to variations among the target group’s characteristics, group sizes, practiced numerical content and the instructional components, and the interventions’ intensiveness in previous studies (see Table 1). In order to further develop the use of computers in learning, it is important to establish the specific components needed for effective, tailored intervention.

The core components of effective numerical intervention for children with (or at-risk for) MD include explicit instructions; repetitive training in basic concepts; step-by-step proceeding; early success; immediate corrective, continuous, and cumulative feedback; and a motivating environment with which to maintain task-orientation (Baker, Gersten, & Lee, 2002; Fuchs et al., 2008; Gersten et al, 2009). Basic skills should be addressed before more complex ones (Dowker, 2001; Sarama & Clements, 2009), and sub-skills should be integrated rather than addressed separately (Fuchs et al., 2012). Further, the relationship between non-symbolic and symbolic notations should be emphasized (Griffin, 2004; Van Luit & Schopman, 2000). However, it seems that previous CAIs for arithmetic have mainly consisted of drill-based practice used for automatizing fact retrieval (see Table 1). Despite generally positive results, the stability of improvements stemming from the intervention is rarely reported. Nonetheless, Kucian et al. (2011) and Shoppek and Tulis (2010) have demonstrated the delayed effects of a short-term intervention that utilized a variety of mathematics content (see descriptions in Table 1).

There is no pre-existing evidence concerning the effects of gradually enhancing children's development of early number skills on learning arithmetic. In addition to group-level intervention effects, we wanted to evaluate individual performance during practice. Therefore, the study goals were: 1) To investigate the group-level effects of short and intensive intervention with a GraphoGame Math program (GGM; see description in section 2.4.) and to assess the stability of any effects; 2) To contrast the intervention participants' gain scores with the performance level of a wait-list control group; and 3) To evaluate the responsiveness to GGM intervention at an individual level by analyzing the game-log data.

Insert Table 1 here.

## 2. Method

Kindergarten teachers from 24 different day care centers in southern Finland were each asked to nominate two children from their group as candidates most in need of extra mathematics intervention. With their parents' permission, the children ($n$ = 48) participated in assessments and computer-assisted intervention. All *48* children were individually tested to determine their early number skills. For ethical reasons, all the children were included in the intervention, even though the assessments indicated that the group of candidates included false positives. Participants were randomized into either 1) a GGM group ($n$ = 24) that took part in GraphoGame Math practice during the first intervention period and had no practice during the second intervention period; or 2) a control group ($n$ = 24) that had no extra practice during the first intervention period and participated in another numerical practice during the second intervention period. All participants were native speakers of Finnish.

### *2.1. Study participants*

Of the *48* participants, *21* children fulfilled the criteria for poor addition skills (1.5 *SD* below the normative age level; below 7th percentiles) and were thus included in the analyses. The inclusion criterion was being unable to solve more than three simple addition problems (e.g., 2 + 1, 1 + 3, 3 + 2) in a time-limited task, as compared to the age-level mean score of 18 out of 45. The reference data had been collected by research assistants for another study one month prior to when the data collection for the current study was carried out (*reference sample n* = 77; *mean age* = 74.2 months, *SD* = 3.6). A sub-sample of *13* children (*4* boys and *9* girls; each from different day care centers) formed the GGM group (*mean age* = 78.6 months, *SD* = 5.4). One participant was not present for the third assessment due to illness. The missing data point was replaced by adding the mean gain score of the GGM group (the difference between the third and second assessments) to the child's score in the second assessment. This did not have an effect on the statistical significance of the findings. A sub-sample of *8* children (*5* boys and *3* girls; each from different day care centers) formed the control (CTR) group (*mean age* = 76.6 months, *SD* = 4.3).

### *2.2. Design*

The study was carried out in the day care centers for 14 weeks, from November to February, including four non-computerized assessment points and two cycles of intervention. After the first assessment (Time 1; week 1), the participants were randomly divided into two groups (week 2). During weeks 3−5, the GGM group played GraphoGame Math, while the CTR group received no extra intervention. After this 3-week cycle, the second assessment was conducted (Time 2; week 6). Next followed a break (weeks 7−9) due to the holiday season. The third assessment was administered after the break (Time 3; week 10). During weeks 11−13, the GGM group received no extra intervention, while the CTR group was offered another type of numerical intervention. Finally, the fourth assessment (Time 4; week 14) was conducted. Due to the different character of the numerical intervention for the CTR group, comparisons are only made for the gain scores until the third assessment.

During the study, all participants followed their normal kindergarten curriculum. According to the Finnish National Board of Education (2010; downloadable in English), an understanding of concepts, classification, comparison, and sorting are specified as objectives for early mathematics in pre-primary education (pp. 11−12).

### *2.3. Early number skill measures*

To assess the reliability the measures of early number skills were piloted before the study. The pilot study sample consisted of kindergarten children from southern Finland ($n$ = 34). Pearson correlation coefficients for test–retest results (over a week) are reported here separately for each measure.

#### *2.3.1. Enumeration*

The enumeration task consisted of a number concept task involving 18 large wooden beads (Salminen, Räsänen, Koponen, & Aunio, 2008a). The child was asked to pick up the number of beads that the tester had requested and then put them on the table. Six different quantities were requested: 3, 6, 8, 10, 13, and 17. The score was the number of correct answers (maximum of 6 points). The Pearson correlation coefficient for test–retest in the pilot study was .72, $p < .001$ (2-tailed).

#### *2.3.2. Verbal counting*

Three different verbal counting tasks were adapted from Diagnostic Tests 3 (Salonen et al., 1994). In the first subtest, the child was asked to count forward starting from 1. A correct response containing number words between 2 and 9 was scored as 1 point, 10−19 as 2 points, 20−29 as 3 points, and 30 as 4 points. In the second subtest, the child counted 4 steps forward from a given number word: 3, 8, 12, and 19. In the third subtest, the child was asked to count 4 steps backwards from a given number word: 4, 8, 12, and 23. In the two latter subtests one point was given for each correctly performed number word sequence. A sum score of the three subtests was used in the analyses (maximum of 12 points). The Pearson correlation coefficient for test–retest in the pilot study was .80, $p < .001$ (2-tailed).

*2.3.3. Number Sets Test*

The original Number Sets Test (Geary, Bailey, & Hoard, 2009) consists of four different parts. In this paper and pencil test, the child was asked to determine as quickly and accurately as possible whether pairs or trios of object sets and/or Arabic numbers match a standard number (5 and 9). Part A involves objects, Part B objects and numbers. Before starting the test, the target number 4 was used when explaining the task. The target number 3 was administered as practice. In this study, only the target number 5 was used. The child was asked to circle the correct pairs or trios that matched 5. A stopwatch was used to measure the time limit of 1 minute per part. The sensitivity for identifying matching pairs or trios (correct responses minus incorrect ones) was scored for both parts, although only Part A was used for the analyses because of a floor effect in Part B. The maximum score was 18 points. The Pearson correlation coefficient for test–retest in the pilot study was .82, $p < .001$ (2-tailed).

*2.3.4. Basic addition*

The basic addition task included single-digit addition problems (a + b = __) in which the sum was 10 or less (Salminen, Räsänen, Koponen, & Aunio, 2008b). The items were pseudo-randomly ordered. Problems in which the sum was five or less were presented first, while highly similar problems were not permitted to follow each other (e.g., 2 + 2 = __ and 2 + 3 = __). Add-zero problems were excluded. The task 1 + 1 = __ was administered as practice, while the tasks 3 + 3 = __ and 1 + 5 =__ (presented vertically) were used for demonstrating a vertical list. The child was instructed to work through the item list as quickly as possible and to respond orally. The time limit of 3 minutes was measured with a stopwatch. The score was the sum of the correct responses (maximum of 45 points). The Pearson correlation coefficient for test–retest in the pilot study was .90, $p < .001$ (2-tailed).

*2.4. Intervention*

The original version of GraphoGame Math (GGM) was designed as part of the GraphoGame project at the University of Jyväskylä in Finland (see Richardson & Lyytinen, 2014; updated version in Finnish and Swedish at www.lukimat.fi). GGM has been modified and further developed as a tool for individual intervention in early number skills for children aged 6 and 7 who are at-risk for MD.

GGM is theory-based and follows the hierarchy of developmental steps in early arithmetic development (e.g., Geary, 2013). The task types are created on the basis of findings from longitudinal studies. To decrease the risk of MD, one-to-one correspondence, comparing, ordering, and object counting should already be emphasized and strengthened at kindergarten age (Desoete & Grégoire, 2006; Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006; Reeve, Reynolds, Humberstone, & Butterworth, 2012). Both number concept (Geary, 2011a; McClelland, Acock, & Morrison, 2006) and counting skills (Aunio & Niemivirta, 2010; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Koponen, Salmi, Eklund, & Aro, 2013) are also included in GGM. The evidence-based intervention principles and components recommended for children with (or at-risk for) MD (e.g., Dowker, 2001; Fuchs et al., 2008; Gersten et al, 2009; Kucian et al., 2011) as well as a personalized user interface with explicit auditory instructions, non-distracting graphics and sound, non-animated stimuli, and corrective and supportive feedback (Seo & Woo, 2010) are also embedded into GGM.

The adaptation in GGM is based on the item difficulty being contingent on the child's performance. The numerical notation proceeds from concrete (objects) to semi-concrete (dots), and finally, to abstract (number words, symbols). Correspondingly, the number range changes from the initial 0−5 to 5−10. An accuracy rate of 85% is required in order to pass a level.

GGM includes six warm-up levels to familiarize the child with the game. After that, one-to-one correspondence (levels 1−3), approximate comparison (levels 4−5), ordering (levels 6−9), and number word−quantity mapping (levels 10−11) are trained. Next, object counting (12−13), exact comparison (14−17), composing (18−19), and decomposing (20−21) are practiced. Finally, conceptual (22−23) and mental addition (24) then symbolic addition (25−) are practiced. While basic addition is the focus after level

21, GGM provides periodic repetition of all the aforementioned basic number content through levels 25−106. Each level consists of approximately 15 trials.

GGM includes two types of trials: either the child selects (with a mouse) the correct target stimulus (e.g., the number 4) among the alternatives to match an auditory cue (e.g., “four”), or the child organizes (with a mouse) randomly ordered stimuli (e.g., four cards depicting varying numbers of beavers) to match an auditory cue (e.g., “Order, the smallest amount first.”). The child clicks the stimuli one-by-one from the bottom of the screen to fill four empty places at the top of the screen. In each trial, the child confirms their response by clicking the “check” button. In the basic addition tasks, the child selects a correct addend/sum to match an auditory cue (selecting trial) or completes/formulates different types of calculations (cf. Riley, Greeno, & Heller, 1983) by organizing addends/sum and/or the arithmetical symbols plus/equal (organizing trial). The two trial types are present throughout the game, but are never featured together within a single level. In GGM, all stimuli remain on the screen until the child clicks on them. There is no time pressure to make the selections. If the child does not click on any stimulus within 30 seconds, the game pauses. These pauses are excluded from the active training time.

GGM provides immediate corrective, continuous, and cumulative feedback. After a correct response, the child hears an expression of approval (e.g., “Great,” or “Well done!”), and the correct stimulus is highlighted in green, while the incorrect stimuli disappear. As corrective feedback after an incorrect response, the child hears a supportive expression (e.g., “Not exactly,” or “Try again”), while the incorrect stimulus is highlighted in red and becomes unselectable. The child selects again until the correct stimulus is found and highlighted in green. After three incorrect selections, GGM provides the correct solution. During each level, continuous feedback is offered in the form of a constantly updated bar at the bottom of the screen showing the proportion of correct responses. After each level, cumulative feedback is offered in the form of stickers available for the child to choose from. The more correct responses the child gives, the more stickers are available. These stickers are collected in a personal, virtual sticker album. This activity is excluded from the active training time.

### *2.5. Procedure*

Kindergarten teachers, trained in the assessment procedure, assessed the children individually at each of the four time points. The assessment tasks were presented in a fixed order: enumeration, verbal counting, Number Sets Test, and basic addition. The assessment sessions took place in a quiet, private room at each child's day care center and lasted approximately 20 minutes. During the 3-week intervention periods, the teachers were asked to arrange a total of between 12 to 15 individual practice sessions during normal kindergarten hours, with 4 to 5 sessions per week. Each session lasted 10 to 15 minutes. The teachers assisted the children with the headphones (equalizing the audio levels) and helped them to start and end their sessions. Individual game-log data on intervention fidelity and performance during practice was collected throughout the study.

*2.6. Data-analysis*

Non-parametric tests were used in the analysis (with SPSS version 20) because the distributions were positively skewed and the sample sizes were small. The Wilcoxon signed-rank test was used for analyzing the within-group (GGM) effects. To calculate the within-group effect sizes for the Wilcoxon signed-rank test results, the following formula was used: ES ($r$) = Wilcoxon $Z/\sqrt{N}$, where $N$ is the number of observations (Field, 2013). The Mann-Whitney U -test was used to analyze the between-group differences (GGM/CTR) at the initial level, as well as to compare the intervention gain scores. Further, to calculate the between-group effect sizes for the Mann-Whitney U -test results, the aforementioned formula was used. The within-group results and the between-group results in gain scores are interpreted with exact, one-tailed $p$ values. The results for the initial level group comparisons are interpreted with exact, two-tailed $p$ values.

## 3. Results

The results are presented in three parts. First, the group-level effects are reported (*GGM* = 13). Second, the between-group differences for the GGM and the CTR groups in terms of gain scores are presented. Finally, the descriptive statistics for individual variance in responsiveness to GGM intervention, as well as individual performance based on game-log data are presented. The means of the raw scores, standard deviations, and medians of the two groups at different time points, as well as the statistically significant intervention effects with the gain scores are shown in Table 2.

Insert Table 2 here.

*3.1. The within-group effects of GraphoGame Math intervention*

The Wilcoxon test showed a significant group-level improvement in basic addition, verbal counting skills, and the Number Sets Test (Part A) between the first and second assessments ($Z$ = -2.99, $p$ = .001, $r$ = .59; $Z$ = -2.87, $p$ = .002, $r$ = .56; $Z$ = -2.29, $p$ = .021, $r$ = .45, respectively). There was no significant immediate effect on enumeration (the scores were close to ceiling; see Table 2).

The Wilcoxon test also revealed that the improvement in basic addition, verbal counting skills, and the Number Sets Test (Part A) remained stable between the first and third assessments ($Z$ = -2.98, $p$ = .001; $Z$ = -2.84, $p$ = .002; $Z$ = -2.36, $p$ = .016, respectively), as well as between the first and fourth assessments ($Z$ = -2.94, $p$ = .001; $Z$ = -2.92, $p$ = .002; $Z$ = -2.29, $p$ = .019, respectively).

*3.2. Intervention effects as compared to the control group*

At the first assessment point, there were no significant differences between the two groups (GGM/CTR) in basic addition, enumeration, or the Number Sets Test (Geary et al., 2009). However, the CTR group had better verbal counting skills than the GGM group (Mann-Whitney $U$ = 18.5, $Z$ = -2.46, $p$ = .012; Table 2).

After the GGM intervention, the differences in gain scores between the GGM and CTR groups did not reach significance (Table 2). Due to the non-significant differences, the between-group comparisons were not analyzed further.

*3.3. Individual variance in responsiveness to intervention*

The children in the GGM group reached the instructed number of intervention sessions ($M$ = 15.31, $SD$ = 1.75, $Mdn$ = 15.00) and minutes (active playing: $M$ = 103.08, $SD$ = 25.87, $Mdn$ = 100.00). However, the game-log data of the *13* children showed a large variance in their performance. The individual progress through the hierarchically organized game content varied greatly, despite initially similar levels of early addition skills. Some of the participants needed a lot of time to complete the easiest non-symbolic contents.

Two figures depict the individual performance during the intervention. Figure 1 depicts the association between the highest played game level (ranging from 12 to 68) and the gain scores for the intervention period in basic addition (ranged from -1 to 15 points). The cases are rank-ordered by the highest played game level they reached in the 3-week intervention period. Level 21 is marked on Figure 1 as a starting point for basic addition contents. In general, it seems that the higher the child progressed in GGM levels, the better the gain in basic addition. This notion receives further support from the significant correlation between the highest level played and the intervention gain scores in basic arithmetic (Spearman's *rho* .80, two-tailed $p$ = .001).

Insert Figure 1 here.

As seen in Figure 1, all the participants passed the first 11 non-symbolic levels (containing one-to-one correspondence, approximate comparison, ordering, and number word–quantity mapping). There was a significant correlation between the rank-orders based on the highest level played and the minutes used for passing the easiest levels

(Spearman's *rho* .85, two-tailed $p < .001$). This reflects the fact that the more time was used for the easiest non-symbolic contents, the fewer levels were passed. The time used for passing the easiest 11 levels is contrasted with the gain scores of the Number Sets Test (non-symbolic task) in Figure 2.

Analysis of the game-log data reveals at least three kinds of profiles. First, there were the children who proceeded rather quickly in both basic number and addition levels, and also improved their skills. Second, there were the children who did not reach the addition levels and who did not improve their addition skills, but who still improved their basic number skills. Third, there were the children who did not seem to respond to the intervention at all, and who improved neither their addition nor basic number skills.

Insert Figure 2 here.

## 4. Discussion

The purpose of this study was to investigate whether intensive computer-assisted intervention (CAI) targeting early number skills and utilizing the recommended principles of effective interventions (Baker et al., 2002; Fuchs et al., 2008; Gersten et al, 2009) and a personalized user interface (Seo & Woo, 2010) would produce positive intervention effects in kindergarteners with poor addition skills. Over the course of a 3-week period, daily practices with GraphoGame Math resulted in within-group improvements in basic addition, verbal counting, and the Number Sets Test. The effect sizes were relatively large (cf. Cohen, 1992), and the intervention improvements remained stable over the 9-week follow-up period. However, there was no between-group difference in gain scores as compared to a wait-list control group.

Previously, stable effects have been reported mainly for teacher-directed training with specific intervention programs (e.g., Dowker, 2001; Griffin, 2004; Van Luit & Schopman, 2000; Wright, 2003), but not after relatively short CAI (e.g., Fuchs et al., 2006). However, utilizing diverse content Shoppek and Tulis (2010; children without mathematics difficulties) and Kucian et al. (2011; children with developmental dyscalculia) reported stable effects in primary school-aged children after intensive CAI. Baroody, Eiland, Purpura, and Reid (2012, 2013) also used mixed training content when introducing CAI to at-risk young children (performance below 25%), but the authors followed the effect only for 2 weeks.

Overall, our results are partially in line with earlier studies in which short, intensive (Cheung & Slavin, 2013) and repetitive training with immediate corrective, continuous, and cumulative feedback has been found to be beneficial for children with (or at-risk for) MD (Baker et al., 2002; Fuchs et al., 2008; Gersten et al., 2009). However, there were no significant between-group differences in terms of intervention gain scores. This could be partly due to the large variance in raw and gain scores revealing individual differences in responsiveness to the intervention. In general, a 3-week intervention period might be too short for children severely at-risk for problems in math development. Also, the sensitivity of assessment tools is often an issue in intervention research. The individuality in responsiveness to intervention, and the attention the children and/or kindergarten teachers paid to the assessments and daily kindergarten activities during the ongoing study, could have affected the variances in outcome measures in such a small scale study. The small sample size both reduces the power for revealing less robust effects, and produces over estimated effect sizes (e.g., Cheung & Slavin, 2013; Slavin & Lake, 2008).

The heterogeneity in responsiveness to the intervention and lack of group-level differences suggest that closer attention should be paid to individual development during the intervention. A dual-discrepancy approach has been suggested for discriminating between responders and non-responders to intervention (poor performance level and intervention growth rate as compared to age-peers; Fuchs & Fuchs, 1998; McMaster, Fuchs, Fuchs, & Compton, 2005). Nonetheless, children with (or at-risk for) MD are not homogeneous in terms of the deficits associated with MD and responsiveness to intervention (e.g., Fuchs et al., 2012; Geary, 2004; Jordan et al., 2003; Rubinstein & Henik, 2009). In order to avoid an inaccurate interpretation of (non-)responding based only on the aforementioned criteria, intervention fidelity and/or outcome measures, performance during intervention should also be assessed in more detail. The utilization of game-log data in CAI provides an opportunity to quantify and evaluate individual development during the intervention. However, it has rarely been utilized in analyses or described (except cf. Obersteiner, Reis, & Ufer, 2013). In this study, the analyses revealed that performance during the intervention varied greatly between participants, despite no problems in intervention fidelity according to both the game-log data and the kindergarten teachers' reports. There were also no marked differences in initial levels of addition skill observed and assessed by the teachers.

Figure 1 shows that the intervention participants who reached the higher game levels generally improved more in basic addition. Due to the hierarchical structure of the game, addition was practiced after the basic levels were satisfactorily passed (after level 21; see Figure 1). Participants who passed the basic levels (except in case number 6) showed larger gain scores for basic addition. It seems logical that the more the child was exposed to basic addition training, the larger the benefit. However, using GGM did not result in an improvement in basic addition skills for all participants.

In sum, it is apparent that children with (or at-risk for) MD would benefit from their progress being monitored (Geary, 2011b). From a pedagogical point of view, in order to detect the non-responders, individual skill levels should be followed both during and after the intervention. In our study, the clearest non-responder was case 9, who used a lot of time to pass the basic levels and who did not improve in any early number skill measures after the intervention or during the follow-up period. Alternatively, some children demonstrated delayed effects in basic addition (cases 6, 12, and 13; Figure 1) and in the Number Sets Test (cases 2, and 12; Figure 2). Case numbers 8 and 10 provide examples of this delayed change in number skill development. To summarize, the performance level of the children who did not respond to the intervention as expected needs monitoring. Non-responders may need more time, or a more varied and individualized intervention, in order to improve their early number skills. Even though the intervention approach used in the current study was based on the theoretical model of early number skill development (cf. Geary, 2013) and the recommended intervention components (cf. Fuchs et al., 2008), it seems that for ensuring progress through the hierarchical content more specific instructional support is needed for overcoming the individual bottlenecks. This could mean a combination of explicit instructions and massed practice with scaffolding cues towards correct responses and sustainable learning.

### *4.1. Limitations and implications for future research*

Even though the assessments included measures with the test-retest reliability information from a pilot study, a baseline would have allowed for a better control for the test-retest effects. When proposing any program as an evidence-based practice for special education, a baseline should be included in order to control for possible confounding factors (Cook & Cook, 2013; Cook, Tankersley, & Landrum, 2009). Obviously, more well-controlled studies of the effectiveness of intensive CAI with larger MD samples are needed, but also, a *triple-discrepancy approach* could be discussed as a way of identifying non-responders (discrepant from age-peers in performance level and intervention growth rate, but also in delayed effect).

### *4.2. Conclusions*

The findings of this study suggest that a carefully planned computerized intervention can support the development of early number skills in some kindergarteners who exhibit poor addition skills. It can provide a tool for close monitoring of the children who progress slowly or who are not responding to intervention, and who are thus in need of more tailored support. Repetition of basic mathematical concepts, proceeding from the concrete to the abstract, and exposure to hierarchically organized content can strengthen a child’s early number skills. However, the individual responsiveness to intervention (Fuchs et al., 2012) needs to be carefully evaluated. It could also be worth to pay more attention to the adaptation and pedagogically meaningful instruction during the response process in order to ensure development during the practice and/or identifying the individual needs of (non-)responders to intervention.

**Acknowledgements and ethical conducts**

The developmental work on GraphoGame Math is carried out as part of the LukiMat -project, which was funded by the Ministry of Education and Culture in Finland (2007−2013). The data used in this study was collected during the same project. The authors declare no potential conflicts of interest with respect to the authorship or publication of this data. The guidelines of the Finnish Advisory Board on Research Integrity were carefully followed in the study. All parents provided written permission for their child to participate in both the assessments and the intervention. All parents were also informed of their right to discontinue participation at any time. Children on the waiting list who were identified as displaying risks in their development of early number skills were also given an opportunity to receive extra support after a short period. In addition, all children in the participating day care centers were allowed to use the intervention programs after the study period if their parents gave consent for them to do so.

**References**

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

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., 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

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*(1), 3–18. doi: 10.1111/j.1469-7610.2005.00374.x

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

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

Christensen, C. A., & Gerber, M. M. (1990). Effectiveness of computerized drill and practice games in teaching basic math facts. *Exceptionality, 1*, 149–165. doi: 10.1080/09362839009524751

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

Cook, B. G., & Cook, S. C. (2013). Unraveling evidence-based practices in special education. The Journal of Special Education, 47(2), 71–82. doi: 10.1177/0022466911420877

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

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

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

Dowker, A. (2001). Numeracy recovery: A pilot scheme for early intervention with young children with numeracy difficulties. *Support for Learning, 16*, 6–10. doi: 10.1111/1467-9604.00178

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., & Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. doi: 10.1037/0012-1649.43.6.1428

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

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., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13*(4), 204−219.

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

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4–15. doi: 10.1177/00222194040370010201

Geary, D. C. (2011a). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539–1552. doi: 10.1037/a0025510

Geary, D. C. (2011b). 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). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science 22*(1), 23–27. doi: 10.1177/0963721412469398

Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The Number Sets Test. *Journal of Psychoeducational Assessment, 27*(3), 265–279. doi: 10.1177/0734282908330592 [Finnish version: Numerosetit (Koponen, T., Salminen, J., Räsänen, P., & Manninen, J. 2008.) Niilo Mäki Institute. Unpublished.]

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

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

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly, 19*, 173–180. doi: 10.1016/j.ecresq.2004.01.012

Hativa, N., & Shorer, D. (1989). Sosioeconomic status, aptitude and gender differences in CAI gains of arithmetic. *Journal of Educational Research, 83*(1) 11–21. Retrieved from http://www.jstor.org/stable/27540363

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74*(3), 834−850. doi: 10.1111/1467-8624.00571

Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*(1), 153–175. doi: 10.1111/j.1467-8624.2006.00862.x

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

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

Kraus, W. H. (1981). Using a computer game to reinforce skills in addition basic facts in second grade. *Journal for Research in Mathematics Education, 12*(2) 152−155. Retrieved from http://www.jstor.org/stable/748709

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönman, C., Plangger, F., … von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage, 57*(3), 782–795. doi: 10.1016/j.neuroimage.2011.01.070.

Käser, T., Baschera, G-M., Kohn, J., Kucian, K., Richtmann, V., Grond, U., … von Aster, M. (2013). Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. *Frontiers in Psychology, 4*, Advance online publication. doi: 10.3389/fpsyg.2013.00489.

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

McMaster, K. N., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*(4), 445−463.

Mevarech, Z. R., & Rich, Y. (1985). Effects of computer-assisted mathematics instruction on disadvantaged pupils' cognitive and affective development. *Journal of Educational Research, 79*(1) 5−11.

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

Okolo, C. M. (1992). The effect of computer-assisted instruction format and initial attitude on the arithmetic facts proficiency and continuing motivation of students with learning disabilities. *Exceptionality, 3*(4), 195−211. doi: 10.1080/09362839209524815

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

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H.P. Ginsburg (Ed.), *The Development of Mathematical Thinking* (pp. 153−196). New York, NY: Academic Press.

Rubinsten, O., & Henik, A. (2009). Developmental dyscalculia: Heterogeneity may not mean different mechanisms. *Trends in Cognitive Science, 13*(2), 92−99. doi: 10.1016/j.tics.2008.11.002

Salminen, J., Räsänen, P., Koponen, T., & Aunio, P. (2008a). The Number Word − Quantity Correspondence Test. Niilo Mäki Institute. Unpublished.

Salminen, J., Räsänen, P., Koponen, T., & Aunio, P. (2008b). The Basic Addition Fluency Test. Niilo Mäki Institute. Unpublished.

Salonen, P., Lepola, J., Vauras, M., Rauhanummi, T., Lehtinen, E., & Kinnunen, R. (1994). *Diagnostiset testit 3. Motivaatio, Metakognitio ja Matematiikka* [Diagnostic Tests 3: Motivation, metacognition and mathematics]. Turku, Finland: Centre for Learning Research, University of Turku.

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

Seo, Y-J. & Woo, H. (2010). The identification, implementation, and evaluation of critical user interface design features of computer-assisted instruction programs in mathematics for students with learning disabilities. *Computers & Education, 55*(1), 363–377. doi: 10.1016/j.compedu.2010.02.002

Schoppek, W. & Tulis, M. (2010). Enhancing arithmetic and word-problem solving skills efficiently by individualized computer-assisted practice. *The Journal of Educational Research, 103*, 239–252. doi: 10.1080/00220670903382962

Shin, N., Norris, C., & Soloway, E. (2006, June 27−July 1). Effects of handheld games on students learning in mathematics. Proceedings of the 7th international conference on learning science (pp. 702−708). Bloomington, IN. Retrieved from http://www.fisme.science.uu.nl/publicaties/literatuur/endnote_ecgbl_472_shin.pdf

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

Trifiletti, J. J., Frith, G. H., & Armstrong, S. (1984). Microcomputers versus resource rooms for LD students: A preliminary investigation of the effects on math skills. *Learning Disability Quarterly, 7*(1), 69−76. Retrieved from http://www.jstor.org/stable/1510263

Van Luit, J. E. H., & Schopman, E. A. M. (2000). Improving early numeracy of young children with special educational needs. *Remedial and Special Education, 21*(1), 27−40. doi: 10.1177/074193250002100105

Von Aster, M. G. (2000). Developmental cognitive neuropsychology of number processing and calculation: varietes of developmental dyscalculia. *European Child & Adolecent Psychiatry, 9*(II), 41−57. doi: 10.1007/s007870070008

Von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49*(11), 868−873. doi: 10.1111/j.1469-8749.2007.00868.x

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: 10.1080/19404150309546674

**Table 1**
Descriptions of the trained contents in computer-assisted intervention studies of basic arithmetic

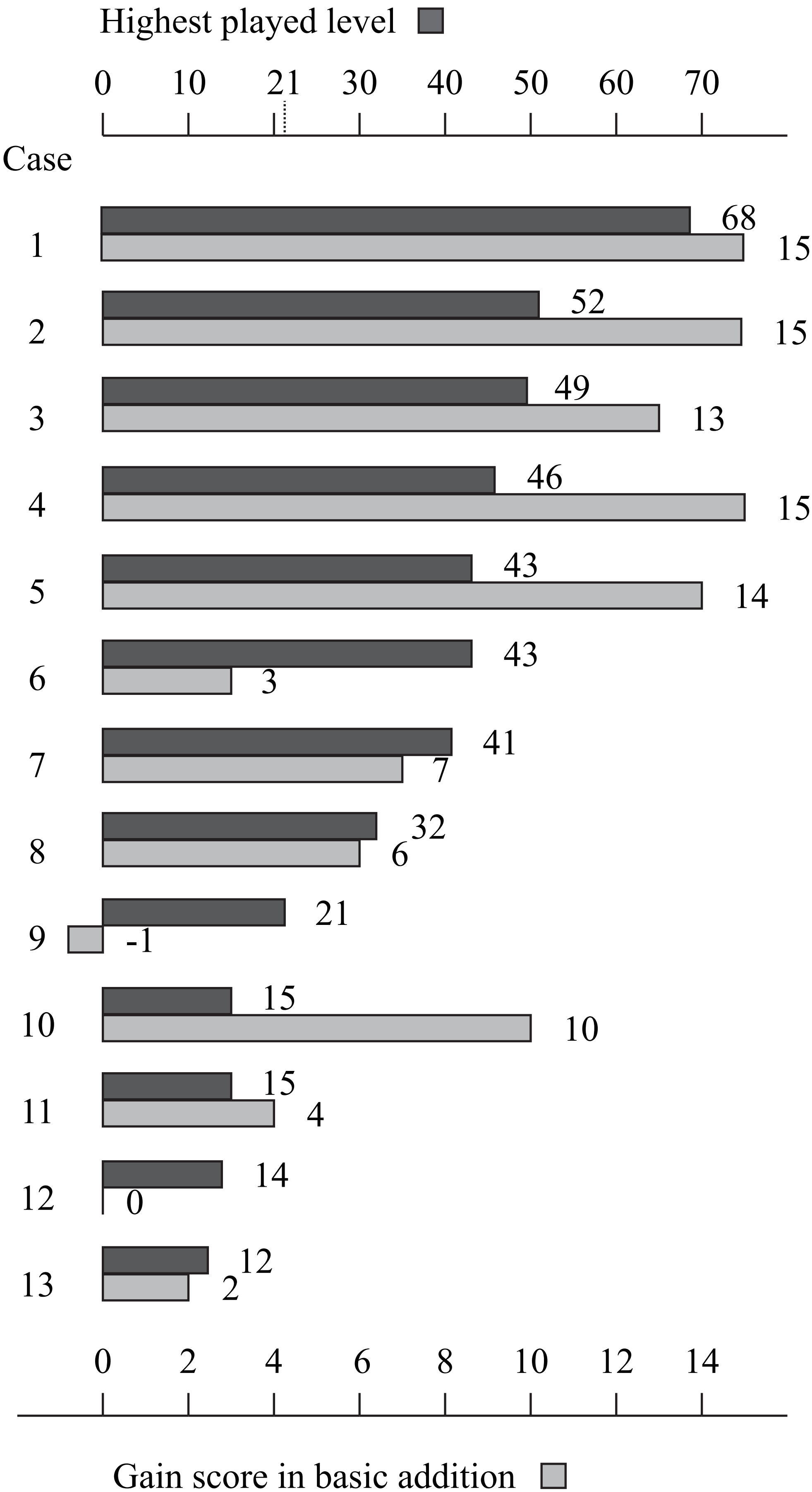
| Study | *n* | Age | Status | Duration | Training sessions | Description of the training | Effect |
|---|---|---|---|---|---|---|---|
| Baroody et al. (2012)[a] | 28 | 5.58 | at-risk[d] | 19 weeks | 20 + 20 x 30 min | Subitizing, enumeration, numeral recognition, transcoding and addition + addition (add-0/1) | + |
| Baroody et al. (2013)[a] | 64 | 6.5 | at-risk[d] | 20 weeks | 20 + 20 x 30 min | Transcoding, verbal counting, object counting, numerical relations, written numbers, arithmetic + addition (add-1/near doubles) | + |
| Christensen & Gerber (1990) | 60 | 8.80 | LD[e] | 2 weeks | 13 x 6 min | Drilling of single-digit addition facts | + |
| Fuchs et al. (2006) | 33 | 6−7 | at-risk MD[d]+RD | 18 weeks | 50 x 10 min | Retrieving addition and subtraction facts | + |
| Hativa & Shorer (1989) | 211 | 8−11 | low SES | semester | 3 times a week | Mixed types of arithmetical contents | - |
| Kraus (1981)[b] | 19 | 7−8 | TA | 2 weeks | 5 + 5 x 15 min | Filling the missing addends to addition combinations | + |
| Kucian et al. (2011)[b] | 32 | 9.5 | DD | 13 weeks | 25 x 15 min | Locating numbers of dots, digits, sums and differences to number line | ++ |
| Käser et al. (2013)[b] | 32 | 7−11 | MD | 12 weeks | 30 x 20 + 30 x 20 min | Number representation, varied types of addition and subtraction tasks | + |
| Mevarech & Rich (1985) | 376 | 8−11 | LA | semester | once a week | Mixed types of arithmetical contents | + |
| Obersteiner et al. (2013) | 147 | 6.91 | TA | 4 weeks | 10 x 30 min | Two versions of Number Race (c.f., Wilson, Dehaene et al. 2006) | + |
| Okolo (1992) | 41 | 9−12 | LD | 9 weeks | 4 x 20 min + 15 min | Mapping presented responds for addition or multiplication facts | + |
| Schoppek & Tulis (2010) | 110 | 8.7 | TA | 10 weeks | 7 x 60 min | Solving arithmetical equations and word problems (addition, subtraction, multiplication, division), number comparison, number line | + |
| Schoppek & Tulis (2010) | 94 | 9.1 | TA | 10 weeks | 7 x 45 min | Described above | ++ |
| Shin et al. (2006) | 46 | 7−8 | middle SES | 18 weeks | 3−4 x 15 min a week | Drilling of addition, subtraction or their mixed combinations | + |
| Trifiletti et al. (1984) | 21 | 9−15 | LD[d] + MD | semester | 40 min a day | Mathematics readiness, addition, subtraction, multiplication, division and fraction | + |
| Wilson et al. (2009)[c] | 53 | 5.6 | low SES | 14 weeks | 6 + 4 x 20 min | Number Race: Approximate comparison between quantities, number symbols and/or addition and subtraction facts | - |
| Wilson, Revkin, et al. (2006) | 9 | 7−9 | LA[f] | 10 weeks | 20 x 30 min | Number Race: Described above | + |

*Note.* Age = Mean age in years (as originally reported). LD = learning difficulties. MD = mathematics difficulties. RD = reading difficulties. SES = socioeconomic status. TA = typically achieving. DD = developmental dyscalculia. LA = low achieving. Effect = Immediate (+) and long-term effects (++) on arithmetic.
[a]Training started with manual games. [b]Training operated at homes. [c]Training mixed with reading software. [d]Cut-off point (not always reported) below 25; [e]16; [f]37 percentiles.

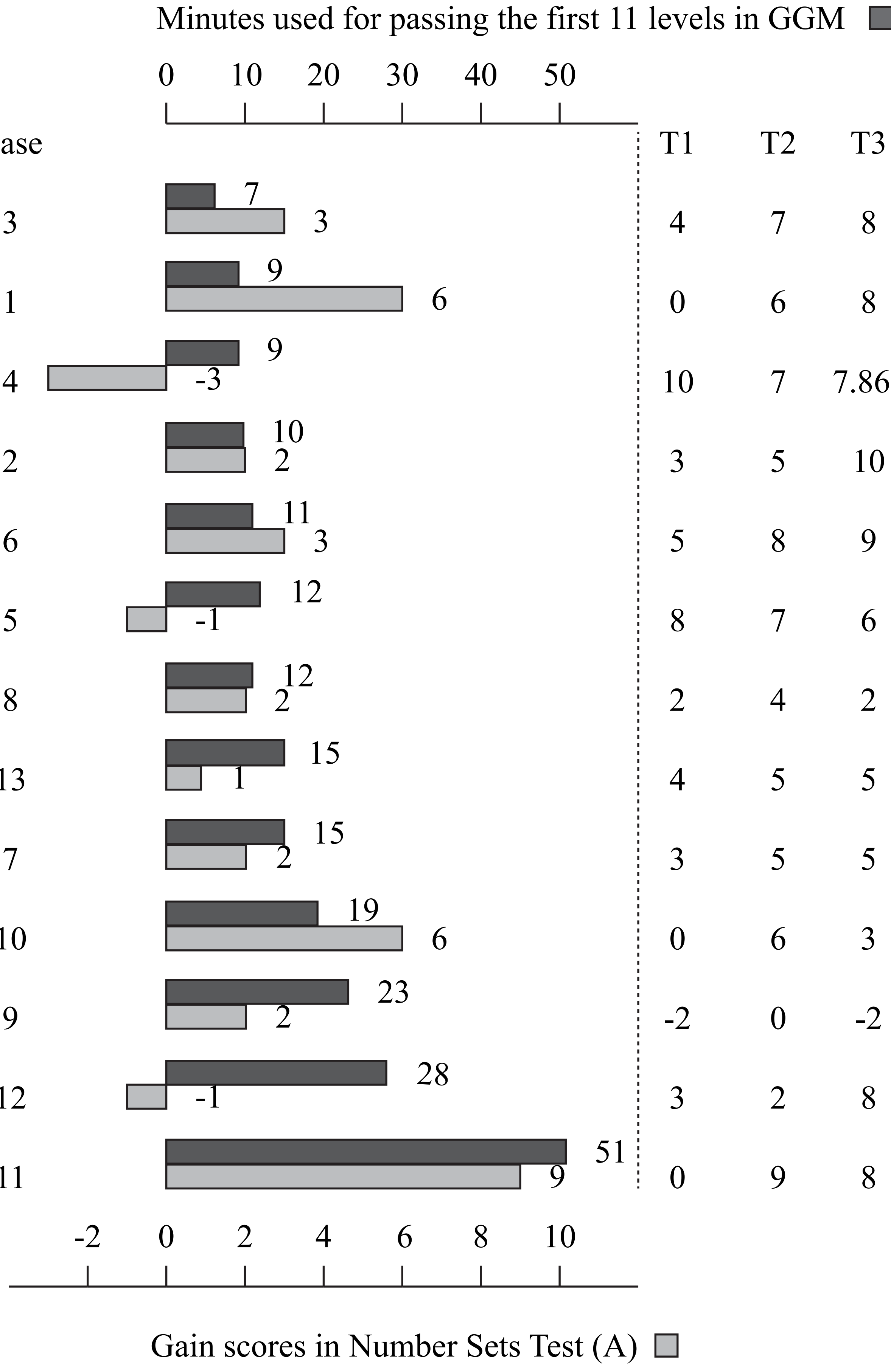
**Table 2**
Group performance scores in different time points and between-group comparisons

| | | Intervention condition | | | | |
|---|---|---|---|---|---|---|
| | | GGM ($n$ = 13) | | CTR ($n$ = 8) | | Group comparisons[a] |
| Variable (max.) | Time | *M (SD)* | *Mdn* | *M (SD)* | *Mdn* | Mann-Whitney U -test |
| Enumeration (6) | 1 | 4.77 (1.09) | 5.00 | 5.25 (0.71) | 5.00 | GGM = CTR |
| | 2 | 4.92 (1.32)[b] | 6.00 | 5.25 (1.04) | 5.50 | |
| | gain 1−2 | 0.15 (1.21) | 0.00 | 0.00 (0.53) | 0.00 | GGM = CTR (.13) |
| | 3 | 5.13 (1.20) | 5.71 | 5.00 (1.51) | 6.00 | |
| | 4 | 5.31 (1.03) | 6.00 | NA | NA | |
| Verbal counting (12) | 1 | **3.92 (1.89)** | **4.00** | 6.88 (2.80) | 5.50 | GGM < CTR* |
| | 2 | **5.54 (2.22)**[c] | **6.00** | 8.00 (2.33) | 8.00 | |
| | gain 1−2 | 1.62 (1.19) | 2.00 | 1.13 (1.55) | 1.00 | GGM = CTR (.14) |
| | 3 | **5.63 (1.80)** | **5.14** | 7.63 (3.25) | 8.50 | |
| | 4 | **7.31 (2.81)** | **8.00** | NA | NA | |
| Number Sets Test (Part A, 18) | 1 | **3.08 (3.33)** | **3.00** | 1.75 (1.83) | 1.50 | GGM = CTR |
| | 2 | **5.46 (2.44)**[d] | **6.00** | 5.88 (5.14) | 4.50 | |
| | gain 1−2 | 2.38 (3.23) | 2.00 | 4.13 (4.79) | 3.50 | GGM = CTR (.18) |
| | 3 | **5.99 (3.36)** | **7.86** | 5.88 (2.30) | 6.50 | |
| | 4 | **5.69 (3.23)** | **6.00** | NA | NA | |
| Basic addition (45) | 1 | **1.08 (1.32)** | **0.00** | 0.38 (1.06) | 0.00 | GGM = CTR |
| | 2 | **9.00 (6.27)**[e] | **10.00** | 5.25 (4.95) | 5.00 | |
| | gain 1−2 | 7.92 (6.05) | 7.00 | 4.88 (5.08) | 3.50 | GGM = CTR (.25) |
| | 3 | **10.18 (7.37)** | **10.00** | 6.25 (4.83) | 6.00 | |
| | 4 | **12.08 (6.79)** | **13.00** | NA | NA | |

*Note.* GGM = Graphogame math; CTR = Performance level control group; NA = non-applicable.
Significant within-group intervention effects are shown in boldface.
[a]Between-group comparisons are made at time point 1, and for gain scores between time points 1−2 (Wilcoxon *ES* ($r$) in parenthesis).
[b]Within-group Wilcoxon *ES* ($r$) = .06 (close to ceiling); [c].56; [d].45; [e].59.
* $p < .05$.




| Case | T1 | T2 | T3 | T4 |
|---|---|---|---|---|
| 1 | 3 | 18 | 24 | 24 |
| 2 | 0 | 15 | 12 | 16 |
| 3 | 3 | 16 | 15 | 15 |
| 4 | 0 | 15 | 17.29[a] | 18 |
| 5 | 0 | 14 | 19 | 20 |
| 6 | 0 | 3 | 12 | 13 |
| 7 | 3 | 10 | 10 | 8 |
| 8 | 1 | 7 | 7 | 1 |
| 9 | 2 | 1 | 0 | 2 |
| 10 | 0 | 10 | 1 | 14 |
| 11 | 2 | 6 | 7 | 7 |
| 12 | 0 | 0 | 1 | 8 |
| 13 | 0 | 2 | 7 | 11 |




| Case | T1 | T2 | T3 | T4 |
|---|---|---|---|---|
| 3 | 4 | 7 | 8 | 1 |
| 1 | 0 | 6 | 8 | 8 |
| 4 | 10 | 7 | 7.86[a] | 11 |
| 2 | 3 | 5 | 10 | 10 |
| 6 | 5 | 8 | 9 | 6 |
| 5 | 8 | 7 | 6 | 7 |
| 8 | 2 | 4 | 2 | 4 |
| 13 | 4 | 5 | 5 | 5 |
| 7 | 3 | 5 | 5 | 6 |
| 10 | 0 | 6 | 3 | 8 |
| 9 | -2 | 0 | -2 | 0 |
| 12 | 3 | 2 | 8 | 3 |
| 11 | 0 | 9 | 8 | 5 |

Figure 1 title:

The association between the highest played GraphoGame Math level and the gain scores of the intervention period in basic addition.

Figure 1 note:

*Figure 1.* Case = intervention participants as a rank-ordered list based on the highest played game level; T1−T4 = individual raw scores in different assessment points in basic addition; Time point 1−Time point 4.
[a]Missing value replaced by adding the mean gain score of the intervention group in basic addition to the case's score in the second assessment.

Figure 2 title:

The association between the time used for passing the basic levels in GraphoGame Math and the gain scores of the intervention period in the Number Sets Test (Part A).

Figure 2 note:

*Figure 2.* Case = intervention participants' case number is based on the highest played game level (presented in Figure 1); T1−T4 = individual raw scores in different assessment points in the Number Sets Test (Geary et al., 2009); Time point 1−Time point 4.
[a]Missing value replaced by adding the mean gain score of the intervention group in the Number Sets Test to the case's score in the second assessment.