European Journal of Special Needs Education

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rejs20

A Mathematics intervention for low-performing Finnish second graders: findings from a pilot study

Riikka Mononen & Pirjo Aunio

Department of Teacher Education, Special Education, University of Helsinki, Helsinki, Finland

Published online: 29 May 2014.

To cite this article: Riikka Mononen & Pirjo Aunio (2014): A Mathematics intervention for low-performing Finnish second graders: findings from a pilot study, European Journal of Special Needs Education, DOI: 10.1080/08856257.2014.922794

To link to this article: http://dx.doi.org/10.1080/08856257.2014.922794

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
A Mathematics intervention for low-performing Finnish second graders: findings from a pilot study

Riikka Mononen* and Pirjo Aunio

Department of Teacher Education, Special Education, University of Helsinki, Helsinki, Finland
(Received 24 March 2014; accepted 6 April 2014)

Evidence-based practice is highly appreciated and demanded in the field of education, especially in relation to extra support provided for children struggling with learning. Currently, there is a lack of intervention studies in the area of mathematics. This study aimed to investigate the effects of a short mathematics intervention programme on second graders (Meanage = 8 years, 2 months) with low performance in mathematics, in a quasi-experimental, intervention-control setting. A group of low-performing second graders (LOWi, n = 11) was taught twice a week for eight weeks on the mathematics topics of number word sequence skills, counting skills and conceptual place value knowledge. The intervention’s effectiveness was examined by comparing the mathematics performance of the LOWi group to those of two groups consisting of low performing (LOWc, n = 13) and typically performing children (TYPc, n = 64), who followed their business-as-usual mathematics instruction. The LOWi group made significant improvements in mathematics but did not show significantly better gains, compared to the LOWc and TYPc groups, immediately and three months after the intervention. We discuss the implications for educational practice and intervention research.

Keywords: intervention; low-performing children; mathematics; second grade

Introduction

Researchers have focused on identifying effective instructional practices in mathematics (e.g. Frye et al. 2013; Gersten et al. 2009; Kroesbergen and Van Luit 2003; Slavin and Lake 2008), but unfortunately, the evidence-based pedagogy is often slow to make its way to classrooms (Forbringer and Fuchs 2014). There is evidence that children with low performance in mathematics may benefit from early pedagogical interventions, which are aimed at improving their mathematics skills and hence trying to diminish the risk of later mathematics difficulties (e.g. Bryant et al. 2011; Jordan et al. 2012; Räsänen et al. 2009). However, the number of evidence-based mathematics intervention programmes in the early school years is still low (Mononen et al. 2014). In this study, we examined a mathematics intervention programme targeting second graders with low performance in mathematics.

Mathematics learning support in the second grade

In grade two, children should master many key mathematics skills that form the foundation for later mathematics learning, such as number sequences forward and...
backwards, enumeration, counting principles, place value knowledge, and addition and subtraction combinations (Sarama and Clements 2009). In Finland, the national core curriculum (NCC) provided by the Finnish National Board of Education (2004) specifies combined aims for first- and second-grade mathematics (for 7–8-year-old children) at a rather general level. Following the NCC guidelines, textbooks with related teacher materials are often used to guide the mathematics teaching in classrooms (Niemi 2010). Typically, these materials for second grade (e.g. Okkonen-Sotka, Sintonen, and Uus-Leponiemi 2009) include topics of teaching addition and subtraction skills with number symbols in the 1–1000 range (first horizontally and then as vertical algorithms), understanding the base 10 and place value system (1–1000), multiplication (with numbers 1–5 and 10), introduction to division and fraction skills with concrete materials, geometry, time and measurement. In grades one and two, three mathematics lessons of 45 min each are provided per week.

Children identified as struggling with the basic mathematics skills (e.g. number sequences, counting and place value knowledge) relevant for later mathematics performance (e.g. Aubrey, Godfrey, and Dahl 2006; Jordan, Glutting, and Ramineni 2010) should be provided with opportunities for more practice of these skills, for instance, in small groups, in addition to core mathematics instruction (Haseler 2008; Riccomini and Smith 2011). In Finland, children are entitled to learning support services based on the three-tiered support model (see Basic Education Act 628/1998 [Finlex 1998]), which shares similarities with the Response to Intervention approach in the US (Gersten et al. 2009). Usually, special education teachers provide part-time support for children needing intensified instruction (i.e. Tier 2 support), either as pull-out lessons or as co-teaching in the classroom. The mathematics learning objectives, teaching methods and materials can then be adjusted according to the children’s needs. At the moment, there is a considerable lack of evidence-based programmes for Finnish children needing intensified support in their mathematics learning.

Results from mathematics intervention meta-analyses have shown the beneficial instructional features of teaching mathematics to low-performing children (e.g. Baker, Gersten, and Lee 2002; Kroesbergen and Van Luit 2003). One of these features is explicit teaching, which includes elements of modelling mathematics concepts and strategies, guided and independent practice opportunities, and continuous feedback (Forbringer and Fuchs 2014). Another proven method to promote learning is the use of visual representations (e.g. cubes, drawings, 10 frames and number lines) of abstract mathematical concepts and strategies, following the sequence of concrete–representational–abstract levels (Witzel, Mink, and Riccomini 2011).

Individual intervention studies for second graders have mainly been conducted in whole classroom settings and have focused on improving two specific mathematics skills: fact fluency (Carr et al. 2011; Duhon, House, and Stinnett 2012; Miller et al. 2011; Poncy, McCallum, and Schmitt 2010) and comprehension of relational terminology within comparison word problems (Schumacher and Fuchs 2012). We found only one intervention study (Bryant et al. 2008) particularly targeting low-performing second graders. In Bryant et al. (2008) intervention study, number concept knowledge (verbal counting, number recognition, comparing and grouping numbers), base 10 and place value understanding, as well as fact knowledge of addition and subtraction, were taught in small groups using explicit instruction. The children practised for 18 weeks, in 62 sessions of 15 min each (a total of 930 min). Using the Texas Early Mathematics Inventories-Progress Monitoring (TEMI-PM)
standardised tests (University of Texas System/Texas Education Agency 2006) as an outcome measure, the study found a positive intervention effect indicated in the addition/subtraction skills subtest \( (b = 0.21, p < 0.05) \). Despite the positive improvement, the low-performing children were unable to close the gap between them and their typically performing peers. Bryant et al. (2008) suggested increasing intervention time so that low-performing children would have more practice opportunities with different representations and more time for fluency building, especially with arithmetic facts.

**Improving mathematics skills in the second grade intervention programme**

Based on the research on the development of mathematics skills, especially from the viewpoint of low-performing children, we developed the *Improving Mathematics Skills in the Second Grade* (IMS-2) intervention programme. It is a small group, intensified intervention programme focusing on practising number sequence skills, counting skills and conceptual place value knowledge in the 1–1000 number range. (See examples of activities in Appendix Table A1.) The IMS-2 programme applies explicit teaching; each lesson consists of a teacher-guided activity to model a new mathematical learning concept or strategy, as well as guided and peer activities (e.g. hands-on activities with manipulatives or card and board games of the current topic), and at the end of the lesson, there is a short, paper-and-pencil individual activity. Mathematical ideas are represented following the concrete–representational–abstract levels, thus giving meaning to abstract concepts using visual representations (e.g. cubes, bundles of sticks, dot cards structured in tens and hundreds). The teacher manual includes 12 lesson plans of 35–45 min each. The lesson plans include specific instructions for teachers to follow in each activity. The manipulatives are made of low cost, everyday materials to be found in every classroom, combined with printable materials (e.g. dot and place value cards) included in the manual. During the developmental work on the intervention materials, two volunteer, special education teachers tried them out with their students and provided practical suggestions to adjust instructions and activities (e.g. to make some instructions easier to understand, to make dot cards bigger, etc.).

**Research question**

This study’s main aim was to investigate the effects of the IMS-2 intervention programme. The following research question was generated: How do low-performing children receiving the IMS-2 intervention develop in their mathematics skills, compared to children in a control group?

**Method**

**Participants**

The current study was part of a larger research project (ThinkMath), which aimed to develop evidence based, small-group intervention programmes in the field of mathematics and thinking skills for 6–8-year-old children. The participants in this study comprised 92 second graders from four classes in schools located in two southern Finnish cities. Only the children who had both the pre- and post-mathematics test
scores were included in the analyses, decreasing the total number by four children \((N = 88; M_{\text{age}} = 8 \text{ years, 2 months}; SD = 3.6 \text{ months}; 48 \text{ boys [54.5%]})\). The attrition rate at the time of the delayed post-test in mathematics was two children \((n = 86)\). All the children had written permission to participate in the study, as authorised by their parents and the municipalities’ educational authorities.

The background information on the children’s home language and their parents’ educational levels was collected by sending a questionnaire to the children’s homes via their teachers. Almost all (97.1%) of the children had Finnish as their mother language. One child spoke Finnish and Swedish, and another spoke Finnish and Spanish. Regarding the highest educational levels of the children’s parents, 5.6% of the mothers had a basic level of education (i.e. nine years of compulsory education), 39.4% had an upper secondary education (i.e. three-year vocational or academic track) and 54.9% had higher education (i.e. university-level degree). For the children’s fathers, the percentages were 14.5, 47.8, and 37.7%, respectively. The statistics for the Finnish population in 2010 showed that 33% of the working population had a basic level of education, 39% had an upper secondary education, and 28% had higher education (Repo 2012). Thus, in our study, the parents had slightly higher educational levels than those of the average Finnish population.

For the research setting, we divided the children into three groups, based on their mathematics performance at the pre-test time. The criterion for low performance was a score at or below the 25th percentile (Fuchs and Fuchs 2006) in the mathematics pre-test (see the test description in the Measures section). Thus, the low-performing intervention group (LOWi) consisted of 11 children (4 boys, 7 girls), the low-performing control group (LOWc) of 13 children (4 boys, 9 girls) and the typically performing control group (TYPc) of 64 children (40 boys, 24 girls).

**Measures**

In addition to mathematics skills, the children’s thinking and language skills were assessed and used as measures of comparability among the groups before the intervention phase.

**Mathematics skills**

A set of core mathematics skills was assessed using the *Assessment of Mathematics Skills in the Second Grade* (AMS-2) (Aunio and Mononen 2012). This group based, paper-and-pencil test includes the tasks of measuring (a) number word sequence skills forward and backwards, (b) numerical relational skills associated to base 10 and place value knowledge, (c) addition and subtraction word problems, and (d) multi-digit addition and subtraction calculations with number symbols, all in the 1–1000 range. Furthermore, single digit (e) addition and (f) subtraction facts in the 1–20 range are measured (each consisting of 40 items and 2 minutes’ time). (See examples of the items in Appendix Table A2.) About 40 min are required to complete the whole test. Each item is scored either zero for a wrong answer or one for the correct answer. We calculated a sum score for scales a–d (a combined scale), which we used to identify the low-performing children in the sample (i.e. scoring at or under the 25th percentile, 0–26 points). Five items (usually the first item of the subscale) did not correlate significantly to the sum score. We did not remove these items, since doing so would not have significantly increased the alpha level.
Cronbach’s alpha was 0.95 for the combined scale (54 items) and 0.94 for both the addition and subtraction facts scales.

Thinking skills
A set of thinking skills was assessed using the Assessment of Thinking Skills in the Second Grade (ATS-2) (Hotulainen, Mononen, and Aunio 2012). This group based, paper-and-pencil test includes tasks measuring two types of inductive reasoning (Klauer 1989): comparing properties and comparing relations. Each item is a multiple-choice task, in which verbal instruction is given. About 40 min are required to complete the whole test. Each item is scored either zero for a wrong answer or one for the correct answer. Ten items did not correlate significantly to the total sum (33 items). Only one item was removed from further analysis, since doing so increased the level of Cronbach’s alpha from 0.68 to 0.70. Hence, after removing one item, the total score for the scale was 32.

Language skills
Reading comprehension and fluency were measured using a standardised reading test for primary grades (Ala-asteen lukutesti [ALLU]) (Lindeman 2005). In the reading comprehension section, a child received a short non-fiction text with 12 multiple-choice questions (four possible answers in each). There is no time limit for the task. Each item is scored either zero for a wrong answer or one for the correct answer. The reliability in terms of Cronbach’s alpha was 0.81. In the reading fluency section, a child has to read four sentences and choose the correct one corresponding to its picture. A two-minute time limit is imposed to answer as many items (maximum of 20) as possible. Since some children accidentally skipped some pages when trying to be quick and therefore did not answer each question in order, we could only count the total score for the test. The reliability in terms of Kuder-Richardson was 0.85, in the study used for norming the test (Lindeman 2005).

Measures of fidelity
The implementation of the intervention has to be closely monitored, using either indirect (e.g. logbooks and teacher reports) or direct (e.g. observations or video recordings) measures or both (Gresham et al. 2000). Implementing the programme as initially intended will enable researchers to understand the relationship between the intervention and the outcome measures (Gersten et al. 2005). We asked the intervention session facilitators to fill out a logbook sheet for every lesson given. They reported the total amount of time (in minutes) used for every lesson and the completed activities. Space was also provided for free comments.

Intervention programme
Improving Mathematics Skills in the Second Grade (IMS-2) (Mononen and Aunio 2012) was used as the intervention programme, which included 12 lessons of 35–45 min each. It focused on practising number word sequence skills, counting and conceptual place value knowledge in the 1–1000 range, following the guidelines of explicit instruction. (See a more detailed description in the Introduction section.)
Procedure
This quasi-experimental study applied a design including the pre-test, intervention phase, post-test, and delayed post-test. The first author and one qualified female teacher with five years of teaching experience conducted all assessment and intervention sessions. The teacher attended a 6-h training before the intervention, together with the other teachers participating in the research project. The training included an introduction to the study procedure and to the assessment and intervention programme. We measured mathematics skills three times: just before the intervention (in September), immediately after the intervention (in December) and three months after the intervention (in March). The thinking and language skills were assessed at the pre-test time. The assessment sessions were held during the school day and took approximately three 45-min sessions on separate days. The intervention sessions were held in small groups of 5–6 children (one group for each facilitator). The initial intention was to have two 45-min intervention sessions per week over six weeks. All the children in the study attended their typical core mathematics instruction. After the delayed measurement time, we conducted a feedback afternoon session for the participating teachers in our research project and presented the results from this and our other similar studies, as well as discussed topics related to the implementation of the intervention.

Data analysis
Due to the small sample size and non-normal distributions found – in the mathematics tests, $D(82) = 0.10$, $p < 0.05$; reading comprehension, $D(82) = 0.15$, $p < 0.05$; and reading fluency, $D(82) = 0.12$, $p < 0.05$ – we used non-parametric tests in all analyses. To examine group differences, we used the Kruskal–Wallis test with post hoc pairwise comparisons, using adjusted significance values. To measure the performance growth within a group, we used a Wilcoxon signed-rank test with the exact method. Effect sizes were calculated as Pearson’s correlation coefficient ($r$) for each gain score comparison among the groups, using the following formula (Rosenthal 1991):

$$r = \frac{z}{\sqrt{N}}$$

where $z$ is the $z$-score value produced from the analysis, and $N$ is the total number of observations. The effect size of $r = 0.10$ can be interpreted as a small effect, $r = 0.30$ as a medium effect and $r = 0.50$ as a large effect (Cohen 1988).

Results
The results are presented in three subsections. First, we analyse the performance before the intervention phase as a whole sample by gender; report the correlations of the parents’ educational levels to the children’s mathematics performance; and compare the mathematics, thinking and language skills performance at pre-test time among the three groups (LOWi, LOWc, and TYPc). Second, we report the intervention effects as a performance growth within each subgroup and compare the gain scores among the subgroups from pre-test to post-test and to delayed post-test time. Third, we report fidelity issues and feedback from the intervention session facilitators.
Performance before the intervention

The Mann-Whitney test showed that boys and girls performed equally well on all pre-test measures (all p-values > 0.05). Both the mothers’ and fathers’ educational levels had a significantly positive correlation (p < 0.05) with the mathematics skills scale (combined scale) (r = 0.243 and r = 0.368, respectively). Further analysis with the Kruskal-Wallis test showed no statistically significant differences (p > 0.05) among the three groups (LOWi, LOWc, and TYPc) concerning the parents’ educational levels.

The means and standard deviations for the mathematics skills measures at each measurement time for the three groups and the pairwise comparisons with significance levels are presented in Table 1. The thinking and language skills measures at pre-test time are shown in Table 2. The Kruskal-Wallis test showed significant differences (p < 0.05) among the three groups in all pre-test measures. Post hoc pairwise comparisons showed no differences between the LOWi and LOWc groups in any pre-test measure. Further investigation revealed however, that in the LOWi group, 10 out of 11 children (90.9%) scored at or under the 15th percentile on the combined mathematics scale, whereas only 5 out of 13 children (38.5%) from the LOWc group did. No statistically significant differences were found between the TYPc and LOWc groups on the scales measuring addition and subtraction facts, thinking and reading comprehension skills. The TYPc group significantly (p < 0.05) outperformed the LOWi group in all measures, except in reading fluency, where no significant difference was found.

Intervention effects

Effects immediately after the intervention

Using a Wilcoxon signed-rank test with the exact method, we found that the LOWi group showed a significant improvement in its performance from pre-test to post-test in all mathematics measures: whole scale, z = -2.94, p = 0.001; addition facts, z = -2.50, p = 0.001; and subtraction facts, z = -2.94, p = 0.010. The LOWc group also made significant improvements in all measures (except for addition facts): whole scale, z = -3.06, p < 0.001; and subtraction facts, z = -2.48, p = 0.011. The TYPc group made significant improvements in all measures: whole scale, z = -6.56, p < 0.001; addition facts, z = -5.95, p < 0.001; and subtraction facts, z = -5.23, p < 0.001.

Gain scores were used in measuring the growth differences among the three groups, since their pre-test scores differed significantly. The gain scores for each mathematics measure were calculated by subtracting the pre-test score from the immediate post-test or delayed post-test score. The Kruskal-Wallis test with post hoc pairwise comparisons was applied. As presented in Table 1, the LOWi and LOWc groups did not differ statistically (p > 0.05) in their gain scores in any mathematics measures. Compared to the TYPc group, the LOWc group had a higher, statistically significant improvement in its combined mathematics score. In the addition and subtraction facts measures, the gains among the three groups did not indicate statistically significant differences.
Table 1. Means, standard deviations and effect sizes for mathematics pre-test, post-test and delayed post-test measures by group

<table>
<thead>
<tr>
<th>Mathematics measure</th>
<th>LOW intervention (1)</th>
<th>LOW control (2)</th>
<th>TYP control (3)</th>
<th>(H(d f = 2)^a)</th>
<th>Pairwise comparisons (b)</th>
<th>Effect size(^c) (r)</th>
<th>Effect size (r)</th>
<th>Effect size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>(M (SD))</td>
<td>(n)</td>
<td>(M (SD))</td>
<td>(n)</td>
<td>(M (SD))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined scale (max. 54 p.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>11</td>
<td>15.36 (5.75)</td>
<td>13</td>
<td>23.69 (2.53)</td>
<td>64</td>
<td>41.69 (7.95)</td>
<td>52.76 *** (3 &gt; 2 = 1)</td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>11</td>
<td>25.45 (11.54)</td>
<td>13</td>
<td>35.69 (7.88)</td>
<td>64</td>
<td>48.66 (4.66)</td>
<td>43.44 *** (3 &gt; 2 = 1)</td>
<td></td>
</tr>
<tr>
<td>Delayed</td>
<td>10</td>
<td>26.10 (10.94)</td>
<td>13</td>
<td>40.54 (7.41)</td>
<td>63</td>
<td>50.32 (4.19)</td>
<td>38.72 *** (3 &gt; 2 = 1)</td>
<td></td>
</tr>
<tr>
<td>Gain pre-post</td>
<td>11</td>
<td>10.09 (7.82)</td>
<td>13</td>
<td>12.00 (6.87)</td>
<td>64</td>
<td>6.97 (5.86)</td>
<td>6.08 ** (2 &gt; 3, 1 = 2, 1 = 3)</td>
<td>-0.16</td>
</tr>
<tr>
<td>Gain pre-delayed</td>
<td>10</td>
<td>11.00 (6.77)</td>
<td>13</td>
<td>16.85 (7.38)</td>
<td>63</td>
<td>8.48 (6.36)</td>
<td>12.40 ** (2 &gt; 3, 1 = 2, 1 = 3)</td>
<td>-0.34</td>
</tr>
<tr>
<td>Addition facts (max. 40 p.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>11</td>
<td>17.45 (4.28)</td>
<td>13</td>
<td>22.85 (6.57)</td>
<td>64</td>
<td>27.61 (6.50)</td>
<td>20.66 *** (3 &gt; 1, 3 = 2, 1 = 2)</td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>11</td>
<td>21.00 (3.55)</td>
<td>13</td>
<td>24.62 (5.39)</td>
<td>64</td>
<td>31.72 (6.23)</td>
<td>30.51 *** (3 &gt; 1 = 2)</td>
<td></td>
</tr>
<tr>
<td>Delayed</td>
<td>10</td>
<td>18.20 (5.05)</td>
<td>13</td>
<td>28.38 (5.39)</td>
<td>63</td>
<td>32.17 (6.82)</td>
<td>23.01 *** (3 &gt; 2 &gt; 1)</td>
<td></td>
</tr>
<tr>
<td>Gain pre-post</td>
<td>11</td>
<td>3.55 (2.95)</td>
<td>13</td>
<td>1.77 (4.89)</td>
<td>64</td>
<td>4.11 (3.74)</td>
<td>2.23 (1 = 2 = 3)</td>
<td>0.18</td>
</tr>
<tr>
<td>Gain pre-delayed</td>
<td>10</td>
<td>1.10 (5.86)</td>
<td>13</td>
<td>5.54 (4.29)</td>
<td>63</td>
<td>4.60 (4.25)</td>
<td>3.69 (1 = 2 = 3)</td>
<td>-0.33</td>
</tr>
<tr>
<td>Subtraction facts (max. 40 p.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>11</td>
<td>11.64 (5.45)</td>
<td>13</td>
<td>17.15 (5.94)</td>
<td>64</td>
<td>21.56 (7.37)</td>
<td>17.96 *** (3 &gt; 1, 3 = 2, 1 = 2)</td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>11</td>
<td>16.00 (4.47)</td>
<td>13</td>
<td>20.69 (6.82)</td>
<td>64</td>
<td>25.22 (6.76)</td>
<td>17.66 *** (3 &gt; 1, 3 = 2, 1 = 2)</td>
<td></td>
</tr>
<tr>
<td>Delayed</td>
<td>10</td>
<td>13.50 (5.21)</td>
<td>13</td>
<td>22.23 (5.88)</td>
<td>63</td>
<td>26.32 (7.14)</td>
<td>21.42 *** (3 &gt; 1, 3 = 2, 1 = 2)</td>
<td></td>
</tr>
<tr>
<td>Gain pre-post</td>
<td>11</td>
<td>4.36 (2.50)</td>
<td>13</td>
<td>3.54 (4.03)</td>
<td>64</td>
<td>3.66 (4.92)</td>
<td>0.76 (1 = 2 = 3)</td>
<td>0.17</td>
</tr>
<tr>
<td>Gain pre-delayed</td>
<td>10</td>
<td>2.60 (4.45)</td>
<td>13</td>
<td>5.08 (5.02)</td>
<td>63</td>
<td>4.79 (5.19)</td>
<td>1.59 (1 = 2 = 3)</td>
<td>-0.24</td>
</tr>
</tbody>
</table>

Note: LOW = low-performing; TYP = typically performing.

\(^a\)Kruskal–Wallis test; \(^b\)Kruskal–Wallis test post hoc pairwise comparisons: significant differences between groups \((p < 0.05)\) marked as \(>\) (e.g. \(3 > 1\) means TYP control group outperformed intervention group), and no difference between groups marked as \(=\); and \(^c\)Effect size \(r\) calculated as \(z\)-value divided by square rooted number of observations. \(*p < 0.05; **p < 0.01; and ***p < 0.001.\)
Effects three months after the intervention ended

The LOWi group showed significant improvement from the pre-test to the delayed post-test only on the combined mathematics scale, $z = -2.81, p < 0.002$. However, this group’s scores on the combined mathematics scale at the delayed post-test time were almost similar to its post-test time performance. The addition and subtraction facts scores of the LOWi children had decreased after the intervention had finished, as shown in Figure 1. In contrast to the LOWi group, the LOWc and TYPc groups showed significant improvements from the pre-test to the delayed post-test in all measures (LOWc: whole scale, $z = -3.18, p < 0.001$; addition facts, $z = -3.09, p = 0.001$; and subtraction facts, $z = -2.75, p = 0.003$; and TYPc: whole scale, $z = -6.73, p < 0.001$; addition facts, $z = -5.99, p < 0.001$; and subtraction facts, $z = -5.74, p < 0.001$). The only statistically significant difference in gain scores among the three groups was found between the LOWc and TYPc groups; accordingly, the LOWc group demonstrated higher improvements in its combined mathematics score from the pre-test to the delayed post-test.

Intervention fidelity and feedback from facilitators

During this study, all the children received their typical core mathematics instruction of three to four, 45-min lessons per week. The facilitators implemented supplemental intervention sessions once or twice a week over eight weeks. All 12 lessons were conducted. One lesson took approximately 42 min, adding up to a total time of 504 min for the intervention. Due to the school timetables, about half of the lessons were given during the school day and the other half as supplemental lessons before or after the school day. The logbook, as well as the information from the intervention facilitators, revealed the children’s difficulties in keeping up with the pace of the intervention programme. Particularly, going over tens and hundreds in a number sequence and adding or subtracting tens to or from two-digit numbers (e.g. add 10 to 34) were difficult for the children. They would have needed more time for practising skills in the 1–100 range before advancing to working on the range beyond 100. According to the facilitators, most of the children showed interest and engagement, especially in game-like mathematics activities. After the intervention, a feedback session with the facilitators revealed that some of the children in the LOWc group had also received some supplemental support for mathematics learning from their school’s special education teacher, although not initially intended. This instruction had concentrated on reaffirming key numeracy skills in the 1–100 range.

Discussion

This study aimed to investigate the effects of the IMS-2 mathematics intervention on low-performing second graders. In two months, those receiving intervention improved significantly in their mathematics skills, but their improvement did not differ from that of the controls, as measured in the gain scores. Three months after the intensified support finished, the mathematics performance of the LOWi group was almost at a similar level, compared to its post-test time performance, but decreased in addition and subtraction fluency.
Regardless of the significant improvements in its mathematics skills from the pre-test to the post-test, the LOWi group did not show better gain scores of statistical significance on the combined mathematics scale, compared to the LOWc group. The reason for this finding may be that some children in the LOWc group also received additional support in mathematics, although not initially intended. The LOWc group even showed a higher improvement score than that of the TYPc group on the combined mathematics scale. Generally, this result would indicate the beneficial effects of additional support in mathematics for low-performing children. However, neither of the LOW groups was able to close the performance gap between them and their typically performing peers on the combined mathematics scale. This finding is similar to those of previous research (e.g. Bryant et al. 2008). The feedback from logbooks revealed the children’s difficulties in keeping up with the pace of the intervention. Dowker (2004) emphasised that mathematics intervention should target the individual needs of the children, thus concentrating on a child’s specific problems. The IMS-2 mathematics programme did not meet all the needs of the children in the LOWi group, and some of them probably would have needed practice in basic skills (e.g. number sequence and counting skills) in the 1–100 range, and a longer intervention time should have been provided (e.g. Bryant et al. 2008).

Three months after the intervention ended, the mathematics performance (measured as a combined scale) of the LOWi group was almost at the same level as its post-test time performance, yet both addition and subtraction fluency had become weaker from the post-test to the delayed post-test time. In contrast, the LOWc group showed continuous improvement in all mathematics skills. One explanation might be that the LOWi group showed more severe difficulties in its mathematics skills, compared to the LOWc group; 90.9% of the LOWi group scored at or under the 15th percentile on the combined mathematics scale, whereas only 38.5% of the LOWc group did. Another reason for the LOWi group’s weaker addition and subtraction fluency at the time of the delayed post-test, compared to the post-test time, could partly be that after the intervention, the children had been focusing mainly on learning multiplication tables, thus not practising their addition and subtraction

### Table 2. Means and standard deviations for thinking and language skills pre-test measures by group.

<table>
<thead>
<tr>
<th>Measure</th>
<th>LOW intervention (1)</th>
<th>LOW control (2)</th>
<th>TYP control (3)</th>
<th>H(df = 2)</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thinking skills (max. 32 p.)</td>
<td>Pre 11 16.82 (5.29)</td>
<td>13 20.08 (2.97)</td>
<td>62 21.95 (4.15)</td>
<td>12.00**</td>
<td>3 &gt; 1, 3 = 2, 1 = 2</td>
</tr>
<tr>
<td>Reading comprehension (max. 12 p.)</td>
<td>Pre 10 5.50 (2.59)</td>
<td>13 7.31 (3.15)</td>
<td>63 8.80 (3.10)</td>
<td>11.23**</td>
<td>3 &gt; 1, 3 = 2, 1 = 2</td>
</tr>
<tr>
<td>Reading fluency (max. 20 p.)</td>
<td>Pre 11 8.22 (3.38)</td>
<td>13 7.69 (2.39)</td>
<td>62 10.63 (3.52)</td>
<td>9.13*</td>
<td>3 &gt; 2, 1 = 2, 1 = 3</td>
</tr>
</tbody>
</table>

Note: LOW = low-performing; TYP = typically performing.  
*Kruskal–Wallis test; **Kruskal–Wallis test post hoc pairwise comparisons: significant differences between groups (p < 0.05) marked as > (e.g. 3 > 1 means TYP control group outperformed intervention group), and no difference between groups marked as =. *p < 0.05; **p < 0.01.
Figure 1. Pre-test, post-test, and delayed post-test performances by group on the mathematics combined scale and addition and subtraction facts measures.
skills. It seems that this group would need frequent reviews of already practised topics (Forbringer and Fuchs 2014) and continuous support supplementing core instruction, in order to retain previously mastered content and prevent widening the achievement gap between them and their peers later on (Aunola et al. 2004; Jordan et al. 2009).

Limitations and future directions

With careful planning before the actual intervention implementation, the researchers can consider a lot of issues affecting the quality of the study (Pruitt and Privette 2001). However, conducting intervention studies in the real school context is challenging. The outcomes were possibly affected by the unexpected pitfalls that we faced in conducting the study (e.g. the LOWc group also received some kind of intensified instruction, although not intended); some drawbacks were a result of our restricted resources (e.g. unable to randomly assign the children into groups) or due to the intervention programme content, which did not adequately meet the needs of the low-performing children in this study (e.g. including too difficult activities for the children).

First, there is a need to further develop our intervention programme in terms of its intensity and content. Previous research has shown that increasing the instruction time (duration in weeks, as well as minutes per session) has led to more positive improvements (Bryant et al. 2008). Increasing the number of sessions to be held over a longer period of time might be more beneficial and provide the children with continuous support throughout the school year. Regarding the programme content, first, working on the 1–100 range should be emphasised, since the children were unable to keep up with the pace of the new concepts and activities in the intervention programme, as reported by the facilitators. Second, the number of game-like activities could be increased in the programme, because the facilitators indicated the children’s interest and engagement in these activities. Previous research (e.g. Sideridis and Tsorbatzoudis 2003) has found that children with learning difficulties are on average, more avoidance oriented and generally lack motivation in learning, compared to their typically developing peers. Therefore, future research might investigate whether using game-like activities as part of the intervention programme provides means for increasing the motivation of low-performing children in their mathematics learning.

Due to our restricted resources, one major limitation in our study was the small number of children involved in the intervention. Although large-scale experimental studies are preferred over quasi-experimental types (Slavin 2008), recruiting teachers and targeting students (especially a specific population, e.g. children with learning difficulties) and randomly assigning the students into groups that receive intervention or not, can be challenging. We were required to divide the low-performing children into intervention and control groups, considering the human resources and practical issues, such as classroom timetables. It is noteworthy whether the teachers should conduct the instruction themselves or the research personnel should do it. The more integrated the instruction is to the normal school work routine and the more it is used with school resources, the more ecologically valid it is (Reed et al. 2013). Volunteer teachers are often highly motivated to go through the whole study process, but this may also affect the outcome of the students’ performance to be too positive (Slavin 2008).
Our limited resources also affected the fidelity issues. In future studies, observations or video recordings in classrooms should be added to the use of logbooks (Gresham et al. 2000) to provide more reliable information on whether the teachers implement the programme as intended, as well as how the children behave and respond to the instruction. Furthermore, more emphasis should be placed on what kind of instruction the control children receive during the intervention phase. The participating teachers should be informed more clearly about the research procedure itself and the expectations for them during the study (Pruitt and Privette 2001).

Conclusion

Evidence-based practice is highly appreciated and demanded in the field of education, but it is often slow to make its way to classrooms (Forbringer and Fuchs 2014). Although researchers would like to provide new, evidence-based intervention tools and methods for schools as quickly as possible, the intervention research process is time and resource consuming. It takes time to develop a programme, conduct intervention studies with rigorous standards and publish the results. Our mathematics intervention for low-performing second graders did not boost their mathematics performance as much as we expected, but this study gave us valuable and encouraging information about the functionality of the programme intensity and content.

Acknowledgements

The research reported here was part of the ThinkMath project, which is supported by the Finnish Ministry of Education and Culture (2011–2015). The opinions expressed are those of the authors and do not represent views of the Finnish Ministry of Education and Culture. We wish to thank all the participating teachers and children.

References


Appendix 1.

Table A1. Mathematics content and examples of activities in the Improving Mathematics Skills in the Second Grade (IMS-2) intervention programme.

<table>
<thead>
<tr>
<th>Lesson number</th>
<th>Content</th>
<th>Example of activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tens: 10 ones equal one 10, tens as quantities and symbols, number sequences by 10</td>
<td>Grouping sticks for bundles of 10; counting bundles of sticks by tens; 10 with place value cards</td>
</tr>
<tr>
<td>2</td>
<td>Tens: 10 tens equal 100, tens and 100 as quantities and symbols, number sequences by 10</td>
<td>Practising tens with dot and place value cards; collect-a-hundred game</td>
</tr>
<tr>
<td>3</td>
<td>Tens: adding and subtracting tens, number sequences by 10</td>
<td>Adding and subtracting tens with bundles of sticks and on a number line</td>
</tr>
<tr>
<td>4</td>
<td>Numbers 10–100: relation of quantities and written numbers, number sequences 0–100</td>
<td>Practising tens and ones with bundles of sticks, and dot and place value cards</td>
</tr>
<tr>
<td>5</td>
<td>Numbers 20–100: counting quantities and writing numbers</td>
<td>Practising writing numbers with dot and place value cards; bingo game</td>
</tr>
<tr>
<td>6</td>
<td>Numbers 20–100: grouping quantities, shortened counting, writing numbers</td>
<td>Counting large amounts of sticks by grouping; introducing ‘Roman numbers’ (X = 10, I = 1) for marking quantities</td>
</tr>
<tr>
<td>7</td>
<td>Hundreds: relation of quantities, said numbers and written numbers</td>
<td>Practising hundreds with dot and place value cards; collect-a-thousand game</td>
</tr>
<tr>
<td>8</td>
<td>Hundreds and tens: relation of quantities, said numbers and written numbers</td>
<td>Practising hundreds and tens with dot and place value cards; car registration plate game</td>
</tr>
<tr>
<td>9</td>
<td>Hundreds and tens: written numbers, going over 100</td>
<td>Building ‘a thousand board’ by tens with place value cards</td>
</tr>
<tr>
<td>10</td>
<td>Numbers 100–1000: constructing numbers of hundreds, tens and ones; relation of quantity and written number</td>
<td>Writing numbers using dot and place value cards; marking quantities by □ = 100, X = 10, I = 1</td>
</tr>
<tr>
<td>11</td>
<td>Numbers 100–1000: number sequences by 1, 2, and 5</td>
<td>Writing missing numbers on a number line; dot-to-dot printables</td>
</tr>
<tr>
<td>12</td>
<td>Revision</td>
<td>Snake-and-ladders game with revision questions</td>
</tr>
</tbody>
</table>
Table A2. Examples of the tasks in the Assessment of Mathematical Skills in the Second Grade (AMS-2).

<table>
<thead>
<tr>
<th>Mathematical skill</th>
<th>Task description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number word sequences (12 items)</td>
<td>Two written numbers are presented. A child must write the next number in the sequence in numeral form (e.g. 50, 51, ___ or 755, 760, ____). A similar task is given for backwards number sequences.</td>
</tr>
<tr>
<td>Base 10 and place value knowledge (24 items)</td>
<td>Number comparison. Four numbers are presented. A child must choose the highest number (e.g. 709, 688, 766, 699)</td>
</tr>
<tr>
<td>Writing numbers. A number is verbally presented. A child must write the corresponding numeral form.</td>
<td></td>
</tr>
<tr>
<td>Adding and subtracting 10 or 100. A number is presented with written instruction to add or subtract 10 or 100 to or from the given number. A child must write the answer in numeral form (e.g. Add 10: 30 ___, 3 ___, 303 ___).</td>
<td></td>
</tr>
<tr>
<td>Addition and subtraction word problems (6 items)</td>
<td>A short addition or subtraction problem is presented verbally. The numbers presented in the problem are written on the board. A child must write the answer in numeral form only (e.g. ‘Jack has 48 candies. Joe has 42 candies. How many more candies does Joe need, to have as many candies as Jack has?’).</td>
</tr>
<tr>
<td>Multi-digit addition and subtraction with number symbols (12 items)</td>
<td>There are 6 addition and 6 subtraction calculations. The calculations include adding/subtracting only ones, only tens or both. Half of the calculations include carrying over 10 or 100.</td>
</tr>
<tr>
<td>Addition facts (40 items)</td>
<td>Forty addition facts (1–20 range) are presented horizontally. Half of the facts include carrying over 10 (the last 20 items). A child has 2 min to write the answers to as many facts as possible.</td>
</tr>
<tr>
<td>Subtraction facts (40 items)</td>
<td>Forty subtraction facts (1–20 range) are presented horizontally. Half of the facts include carrying over 10 (the last 20 items). A child has 2 min to write the answers to as many facts as possible.</td>
</tr>
</tbody>
</table>