Riikka Mononen [1]
Pirjo Aunio [2]
Tuire Koponen [3]
Mikko Aro [4]

A Review of Early Numeracy Interventions for Children at Risk in Mathematics

Abstract

This study reviewed early numeracy interventions for four- to seven-year-old children at risk for mathematics difficulties. The search yielded 19 peer-reviewed studies with pre- and post-treatment control designs. The interventions were categorised as either core or supplemental instruction. The study analysed the effectiveness and identified the pedagogical components of the interventions: setting, duration, numeracy content used for intervention training and progress measurement, conductor and professional developmental support offered, and instructional design features. The interventions showed, to various degrees, the promising effect of improving the early numeracy skills of at-risk children. Results indicated that different types of instructional design features, including explicit instruction, computer-assisted instruction (CAI), game playing, or the use of concrete-representational-abstract levels in representations of math concepts, led to improvements in mathematics performance. The paper discusses the implications for practise and suggestions for future research.

Key Words: *At-risk children, early numeracy, intervention, mathematics, review*

[1]M.Ed., University of Helsinki, Department of Teacher Education, Special Education, PO Box 9, 00014 University of Helsinki, Finland,
e-mail: riikka.m.mononen@helsinki.fi
Tel.: +358 50 318 2383

[2] PhD., University of Helsinki, Department of Teacher Education, Special Education, Finland,

[3] PhD., Niilo Mäki Institute, University of Jyväskylä, Finland,

[4] PhD., University of Jyväskylä, Department of Education, Finland.

Introduction

Studies in the field of early numeracy have reported large individual differences among children before the onset of formal education (Aunio, Hautamäki, Sajaniemi, & Van Luit, 2009; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Wright, 1991). Longitudinal studies have shown that early numeracy skills are accurate predictors of later mathematics achievement (Desoete, Ceulemans, De Weerdt, & Pieters, 2010; Desoete, Stock, Schepens, Baeyens, & Roeyers, 2009; Krajewski & Schneider, 2009; Missall, Mercer, Martínez, & Casebeer, 2012). Moreover, children identified as struggling with early numeracy in kindergarten seem to develop mathematics skills slower than children with no difficulties do (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Morgan, Farkas, & Wu, 2009). Without extra support, intervention, children are likely to remain low-performers throughout their primary school years (Aubrey, Dahl, & Godfrey, 2006). A number of reviews and meta-analyses concerning the effects of mathematics interventions have focused mainly on school-age children with, or at risk for, mathematics difficulties (MD) (Baker, Gersten, & Lee, 2002; Codding, Hilt-Panahon, Panahon, & Benson, 2009; Gersten, Chard et al., 2009; Kroesbergen & Van Luit, 2003; Kunsch, Jitendra, & Sood, 2007). The findings from these studies have guided the teaching of mathematics to these children. So far, only one meta-analysis has addressed preschool and kindergarten children's mathematics learning (Malofeeva, 2005). Malofeeva found early mathematics instruction to be effective (g = .47). However, the types of participants were not controlled. This review aims to fill this research gap, by reviewing studies of interventions designed to enhance early numeracy skills of young children aged 4–7 at risk for MD.

Early Numeracy Learning and Difficulties

During early childhood years, children develop several mathematical skills that form the foundation for later mathematics learning at school (for a detailed description of learning trajectories, see Sarama & Clements, 2009). Counting skills, basic arithmetical skills, magnitude comparison, and numeracy-related logical abilities, measured before the onset of formal schooling have been shown to be good predictors of later mathematics learning in school (e.g., Aubrey et al., 2006; Aunola et al., 2004; Aunio & Niemivirta, 2010). Therefore, it is important that early mathematics instruction emphasizes these skills. According to Malofeeva (2005), the five most trained mathematics skills with preschoolers and kindergartners were counting, conservation, computation, enumeration, and comparison. Measurement, estimation, writing numerals, and fractions were less emphasized.

Literature uses several terms to describe the levels of difficulties in mathematics performance. This reflects the unsystematic use of terminology. Term *developmental dyscalculia* (DD, also referred as mathematics disability) represents the most severe and persistent mathematics difficulties. Accordingly, children with DD show lower mathematics performance than expected for their age (i.e., falling below the 10th percentile on a standardized mathematics test) (Mazzocco, Feigenson, & Halberda, 2011). The term *mathematics difficulties* (MD) is often used with individuals who show performance between the 11–25th percentile (Mazzocco et al., 2011). Some (e.g., Fuchs & Fuchs, 2006) refer this group to as *low-performing* or *low-achieving* in mathematics, suggesting their risk for severe MD. The moderate or mild mathematics difficulties of this group may stem from external factors, such as poor teaching, low socio-economic status, and behavioral attention problems or domain-general cognitive deficits (Price & Ansari, 2013). In this study, we use the term low-performing in mathematics to refer all children who perform below average mathematics skills, thus at risk for mathematics difficulties (mild, moderate, or severe).

Pedagogical Support in Early Numeracy Learning

Numeracy instruction prior to formal schooling should prepare children with adequate skills to follow mathematics instruction at school. Since there is great variation in children's skills before the onset of formal schooling, different levels of interventions are needed in pre-kindergarten and kindergarten to support the diversity of children. Whereas the term *mathematics core instruction* is used in the context of providing instruction for all children in a classroom, the term *mathematics intervention* is most often used in the context of promoting low-performing children's learning with more intensified, supplemental instruction (Riccomini & Smith, 2011; Haseler, 2008). Research suggests that effective early mathematics core instruction can serve as the first approach for improving the mathematics performance of kindergarten children, including those at risk for MD (Clarke, Doabler, et al., 2011). Here, the term *intervention* covers core or supplemental instruction, the latter given usually in small groups or individually.

To promote the early learning of numeracy appropriately, teachers should be familiar with the developmental path of numeracy skills, as well as the problems a child may encounter (Anthony & Walshaw, 2009). Traditionally, pre-kindergarten and kindergarten instruction has been informal, often given during playtime or in games, and additional support may have not been offered to children struggling in mathematics (Clarke et al., 2011). The objectives of early numeracy learning, as set in the kindergarten core curriculum, are usually narrow and less structured than those in grades one and above. Ideally, core instruction is built around the most critical mathematics content, and reflects what is currently known about the instructional design features that research has found to be effective for enhancing the mathematics performance of low-performing children (Clarke, Doabler, et al., 2011; Codding et al., 2009; Gersten, Chard, et al., 2009). Thus, according to current knowledge, good core instruction includes (1) explicit and systematic instruction, which takes the form of modelling or demonstrating how to solve a problem by using specific procedures, breaking tasks into smaller units, providing a cumulative review, and providing error correction procedures (Bryant, Roberts, Bryant, & DiAndreth-Elkins, 2011), and (2) the use of visual representations, such as cubes, drawings, 10 frames, and number lines, in the introduction of mathematics ideas at the concrete-representational-abstract (CRA) levels (Witzel, Mink, & Riccomini, 2011).

For children struggling with key concepts that are deemed important for later achievement in mathematics, supplemental, more intensified instruction should be provided. Typical elements for supplemental instruction listed in the literature (e.g., Gersten, Beckmann et al., 2009) are as follows: (1) increased levels of targeted instruction in specific mathematical skills in addition to core curriculum instruction; (2) instruction in small-group settings or individually outside the whole group by an educator, class aide, or mathematics interventionist; and (3) frequent monitoring of progress in order to determine whether the instruction is effective in improving learning outcomes.

Objectives of the Present Review

The purpose of the present review was to examine the effectiveness and pedagogical implementation of early numeracy interventions (i.e., core and supplemental instruction) for four- to seven-year-old children at risk for MD. The study compared children's performance in intervention groups with children's performance in control groups. Standardized effect sizes (g) were used as a measure of this relationship, as the outcome measures differed across studies. In addition, the study identified common pedagogical components of interventions likely to influence their effectiveness. Therefore, interventions were considered from viewpoints of (a) setting, (b) duration, (c) numeracy content used for intervention training and progress measurement, (d) conductor and professional developmental support offered, and (e) instructional design features. Furthermore, the study examined interventions from the points of view of reducing the performance gap between low-performing children and age-related, typically-performing peers and of determining the delayed effects of learned skills after the intervention had finished.

Method

Literature Search Procedures

To identify a broad range of studies, we searched the recent international literature using several educational databases (ERIC, JSTOR, PsycINFO, EBSCO, and Ovid) with different combinations of key words, such as early number skills, early numeracy skills, kindergarten math(ematics), mathematics education, intervention, mathematical difficulties, mathematical disabilities, counting, comparison, and arithmetic. We then used references within articles to find more studies.

Criteria for Inclusion

The criteria for the inclusion of studies were set tight in order to reduce bias in the results. The following criteria addressed participants, interventions, comparisons, and outcomes, as suggested by O'Connor, Green, and Higgins (2008):

1. The studies evaluated early numeracy intervention programmes for children who were at risk for MD in the age range 4–7 years. The study sample included children performing lower than expected for their age (LOW) in mathematics or having a background of low socio-economic status (SES) and thus at risk for MD (Jordan, Kaplan, Olah, & Locuniak, 2006). The cut-off criterion for low-performing children was specified in the primary study, such as scoring below the 25th percentile on a mathematics test. We excluded studies focusing only on typically performing children (TYP), or children with special educational needs (SEN).

2. The studies used random assignment or quasi-experimental design. Although using only randomized studies in reviews and meta-analyses is recommended, quasi-experimental studies produce good estimates of intervention outcomes as well, if attention is paid to selection bias (Slavin, 2008). Quasi-experimental studies required that the initial pre-test equivalence in mathematics skills between intervention and control groups was indicated.

3. Studies without control groups were excluded. Control groups were (1) active control groups (CA) receiving similar amounts of instruction; for example, regular math classroom activities or another type of intervention, such as reading, (2) passive control groups (CP) with no supplemental instruction provided, or (3) other types of mathematics intervention groups (CI), receiving other mathematics interventions. Comparisons of at-risk children to TYP children were considered additional data to determine if the interventions could reduce the performance gap between these groups, or as a control group (CP), if the study included only mathematics interventions.

4. The sample size was allowed to vary. However, the study exercised caution when interpreting the results and effect sizes of studies with small sample sizes, which tend to have more extreme effect sizes (Slavin, 2008).

5. A teacher or members of the research team, such as a trained tutor, implemented the intervention.

6. The duration of the intervention was allowed to vary. As Slavin (2008) previously suggested the criterion of 12-weeks, we used this as a benchmark for both short and long interventions. We considered an intervention short if the post-test was given within 12 weeks from the beginning of the intervention.

7. Dependent measures included reliable quantitative measures of mathematical performance. In addition, the measure had to be fair for the control groups, and not favour the experimental group.

8. The studies had to provide sufficient data for effect size calculations. Either they presented post-test data as means adjusted for pre-test performance with non-adjusted standard deviations, or they presented both pre- and post-test data as unadjusted means and standard deviations. In addition, if the study did not report such values, but statistical information from *t*-tests or ANOVA was available, it was used if group equivalence existed on pre-test measures achieved through random assignment.

9. The studies could take place in any country, but had to be published in English in a peer-reviewed journal between 2000 and 2012.

Nineteen studies met these criteria. Five were identified as core instruction interventions and fourteen as supplemental instruction interventions. The studies are listed in Appendix A; Table A1 indicates core instruction and Table A2 indicates supplemental instruction. The excluded studies and the reasons for their exclusion are listed in Appendix B.

Coding of Studies

From each included study, the following information was coded: study characteristics (authors, year, and country), methodological characteristics (design, measures, reliability, and control group status), sampling characteristics (number of participants, mean age, and at-risk status: LOW or SES), and components of intervention (programme, duration, setting, leader, professional development provided, fidelity, instructional design features, materials, and practised and measured numeracy content).

Calculation of Effect Sizes

We calculated the effect sizes for each mathematical performance dependent measure using Hedges' *g* with correction for small sample sizes (see Turner & Bernard, 2006). The differences in the outcomes of the experimental and control groups were calculated after adjustment for pre-test differences. Two studies used analysis of covariance (ANCOVA), and in these, Hedges' *g* was calculated as a covariate (pre-test) adjusted mean difference divided by the unadjusted pooled within-group post-test standard deviation (SD) (What Works Clearinghouse, 2008). In the fourteen studies that did not present adjusted means, but presented unadjusted pre-test and post-test means with SDs, we calculated the difference of the mean pre-post change in the experimental group and the mean pre-post change in the control group, divided by the pooled within-group, pre-test SD (Morris, 2008). Three studies did not provide sufficient means or SDs, but had group equivalences on pre-intervention measures achieved through random assignment. With these studies, we used presented *t*-test values to calculate Hedges' *g*. We calculated the confidence intervals (95%) for effect sizes by using the standard error of the effect size estimates (Turner & Bernard, 2006). For the purpose of interpretation, we used Cohen's (1988) distinctions of the magnitude of the effect (small effect 0.20–0.49, moderate effect 0.50–0.79 and large effect 0.80+). Only six studies used more than one post-test, as post-intervention follow-up. To calculate the effect sizes, we used only the scores of the first post-test. We described the delayed effects of interventions, whether the intervention effect faded or held, and reported statistical significances as they were given in the primary studies.

Results

The findings of intervention studies are reported according to study and sample characteristics, setting, duration, numeracy content used for intervention training and progress measurement, conductor and professional developmental support, instructional design features, effectiveness in reducing the performance gap between children at risk for MD and typically developing age peers, and delayed effects of intervention. A summary of intervention descriptions with statistical significances based on the reported values in the primary studies, calculated effects sizes, and delayed effects is included in Appendix A (Tables A1 and A2). The findings of core and supplemental instruction interventions are treated separately.

Core Instruction Interventions

We found five intervention studies identified as core instruction of which two focused on five-year-old LOW children (Chard et al., 2008; Clarke et al., 2011) and three on four-year-old SES children (Clements & Sarama, 2007, 2008; Clements, Sarama, Spitler, Lange, & Wolfe, 2011). No interventions were found for six- or seven-year-old children. All studies were conducted in the USA. Three studies applied randomized designs. A total of 3127 children participated in the studies (n = 1904 in experimental groups, n = 1223 in control groups). Sample sizes in the experimental groups ranged from 30 to 927 children. The interventions focusing on LOW children used the same programme: Early Learning in Mathematics (ELM). The Building Blocks (BB) programme was used with the SES children.

Setting

All core instruction interventions were implemented in whole group settings in classrooms. Every intervention, however, included practising in small groups or with peers within the classroom; however, children did not leave the classroom to work with another teacher. In BB interventions, the children also had access to computer software linked to the learning objectives of the programme.

Duration

Common to all interventions was the long duration, from 25 to 26 weeks. The exact time used for practise varied from 2,470 minutes to 7,200 minutes. Two studies (Clements & Sarama, 2007; Clements et al., 2011) did not specify the frequency and time used apart from daily intervention instruction over a school year. In one BB study (Clements & Sarama, 2008), more instruction time was used among the BB group than among the controls: the CI group used almost 800 minutes less than the BB group, while the CA group's instruction time varied, depending on the school, but was still less than in the

BB group. A comparison of the studies using the ELM programme showed that a longer instruction time, a total of 3,000 mins versus 7,200 mins during the school year, produced a significant difference in the LOW children's mathematical scores compared to the CA group's LOW children participating in typical mathematical activities.

Numeracy content used for intervention training and progress measurement

Both the BB and the ELM programmes included a variety of early mathematical learning objectives for the targeted age group, such as recognition of numbers, object and verbal counting, comparison skills, ordering skills, simple addition and subtraction, place value knowledge, geometry, and measurement. Mathematics knowledge was measured using either standardized or developer-based tests, or both. Tests measured several of the skills trained in the interventions, but did not specifically favour the experimental groups. All tests were adequate in their reliability (i.e., $\alpha > .60$).

Conductor and professional developmental support

Teachers implemented all core instruction interventions, replacing the mathematics instruction typically used in the classroom. Implementation was observed in every study to ensure the fidelity of the instruction and to confirm that programme delivery was as expected. In each study, before and during the interventions, experimental group teachers participated in professional developmental (PD) training sessions, varying from one-half day to thirteen days. In addition to introducing the principles and instructional practises of the programme, some research groups offered general professional development in mathematics teaching, such as information about research-based effective practises in mathematics (Clarke et al., 2011) or learning trajectories in mathematics (Clements et al., 2011). The importance of PD was emphasized as the key part of successful BB intervention (Clements & Sarama, 2008; Clements et al., 2011). In addition to several PD days during the intervention phase, one BB study (Clements et al., 2011) included support from trained mentors familiar with the programme. In contrast to other studies in which control groups received PD support time not similar to the experimental group or did not receive any, one of the two control groups (CI) in the study of Clements and Sarama (2008) received the same amount of PD as the experimental group did. The training included the same topics, but always in the context of the specific curriculum. Regardless, BB showed a significantly better intervention effect than the CI group.

Instructional design features

The ELM programme integrated the use of research-based instructional principles throughout, such as explicit instruction with teaching concepts using CRA levels, mathematics-related discourse, and frequent and cumulative embedded reviews, both within and across the lessons. Hands-on material and worksheets were primarily used. The BB programme shared some features (e.g., explicit instruction) with ELM. Contrary to the ELM programme, the BB programme included more differentiation in classroom work by including small-group activities and individual computer work. Hands-on materials, books, and games were also used in the BB.

Effectiveness in reducing the performance gap and delayed intervention effect

Although at-risk children made significant improvements in their learning, only one study (Clarke et al., 2011) specifically reported the performance of at-risk children compared to their age peers after the intervention. About 42% of LOW children who had been given the ELM intervention were re-classified into the no-risk category after the intervention, compared to about 33% of the LOW CA group's children. None of the studies applied follow-up measurements to demonstrate continuation of the intervention effect post-intervention.

Supplemental Instruction Interventions

Fourteen intervention studies were identified as supplemental instruction. Six were aimed at LOW children (Aunio, Hautamäki, & Van Luit, 2005; Bryant et al., 2011; Fuchs, Fuchs, & Karns, 2001; Fuchs et al., 2005; Fuchs et al., 2006; Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009) and eight focused on low SES children (Arnold, Fisher, Doctoroff, & Dobbs, 2002; Dyson, Jordan, & Glutting, 2011; Jordan, Glutting, Dyson, Hassinger-Das, & Irwin, 2012; Klein, Starkey, Clements, Sarama, & Iyer, 2008; Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Sood & Jitendra, 2011). Twelve studies were conducted in the USA, and two in Finland. All but one applied random designs. Five interventions were aimed at four-year-old children, five at five-year-old children, and four at six-year-old children. The six-year-olds were in the first year of formal schooling (first-grade students in the USA), except in one study (Räsänen et al., 2009) conducted in Finland, where formal schooling begins at the age of seven, and therefore, six-year-old children were kindergartners. No intervention studies were found for seven-year-olds. A total of 2081 children participated in the studies (n = 807 in experimental groups, n = 1274 in control groups). Sample sizes in experimental groups ranged from 15 to 139 children.

Setting

Three interventions used whole-group settings, in which instruction replaced some portion of typical classroom mathematics instruction (Arnold et al., 2002; Fuchs et al., 2001; Sood & Jitendra, 2011). In the classroom, practising could include working in whole and small groups or working in pairs, or individually with a computer. Seven interventions were delivered in small-group settings outside the classrooms (Aunio et al., 2005; Bryant et al., 2011; Dyson et al., 2011; Fuchs et al., 2005, 2006; Jordan et al., 2012; Klein et al., 2008). Two of these studies (Fuchs et al., 2005, 2006) included individual work with computers. Four studies included one-to-one instruction, including playing either a board game (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009) or a numeracy game on a computer (Räsänen et al., 2009).

Duration

Intervention durations varied greatly: from 2 weeks to 36 weeks. Eight studies were short in their duration (*Mdn* = 3.5 weeks, range 2–8 weeks), and six studies were long (*Mdn* = 18.5 weeks, range 15–36 weeks). The median total time in minutes used in short interventions was 225 minutes (range 60–720 minutes), and in long interventions 1,330 minutes (range 500–1,900 minutes).

Numeracy content used for intervention training and progress measurement

We separated supplemental interventions into two categories based on whether the intervention concentrated on practising a variety of mathematical skills or a few specific skills. Among multi-skill interventions (Arnold et al., 2002; Aunio et al., 2005; Bryant et al., 2011; Dyson et al., 2011; Fuchs et al., 2001; Fuchs et al., 2005; Jordan et al., 2012; Klein et al., 2008; Sood & Jitendra, 2011), the six most practised skills were recognition of numbers (100%), object counting (100%), verbal counting (88.9%), comparing (88.9%), ordering (66.7%), and early addition and subtraction (66.7%). The operational numbers in activities were generally within the 1–20 range. The two least practised skills were geometry (22.2%) and estimating (0%). In interventions focusing on specific mathematics skills using computer-assisted instruction (CAI), children practised addition and subtraction facts (Fuchs et al., 2006), and counting and comparing (Räsänen et al., 2009). Playing a linear board game (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009) aimed to promote children's numerical knowledge. Mathematics knowledge with multi-skill interventions was measured using either standardized tests (*n* = 4) or developer-based tests (*n* = 2), or both (*n* = 3). Tests measured a variety of mathematical skills with adequate reliabilities reported. Measures used in specific targeted interventions were all developer-based. Reliabilities for tests were acceptable. Intervention studies using The Great Race board game did not report reliabilities for the measures used.

Conductor and professional developmental support

Teachers implemented five interventions and members of the research team, trained tutors, implemented eight. In one study, both teachers and members of the research team were involved. The teachers were observed in every study to ensure fidelity - that the programme was implemented as intended. In each study, before and during the intervention, experimental group teachers participated in PD training sessions, which took from two to eight hours. The training introduced the principles and instructional practises of the programme. Some researchers also offered informal support during the intervention phase. For example, when a member of the research team visited the classroom, the teacher could ask questions related to the intervention. In the studies using CAI, the conductors did not receive PD training, as they primarily supervised the children working on the computer.

Instructional design features

Explicit instruction occurred in half of the interventions (n = 7). One study combined guided instruction with explicit instruction. One intervention used peer-assisted tutoring, and one incorporated math activities into daily routines. Four studies (28.6%) used CRA in teaching mathematical concepts. Games were used in five studies (35.7%), and in three of these studies, a board game was used as the primary instructional material. In addition to interventions using only CAI (n = 2), two other interventions included practise with mathematics software as part of the instruction. Computer programmes provided corrective feedback for each child's performance.

Effectiveness in reducing the performance gap and delayed intervention effect

Seven studies out of fourteen provided enough information to make comparisons on post-intervention performance between at-risk and TYP children. Although at-risk children made significant improvements in their learning, they were able to reach the performance level of their age peers in only three studies. In the study of Aunio et al. (2005), the LOW children reached the level of TYP controls in counting skills. Bryant et al. (2011) reported that about 45% of LOW children were re-classified into the no-risk category after the intervention, compared to about 22% of the LOW CA group's children. After receiving a short board game intervention (Siegler & Ramani, 2008), the SES children performed at the same level as their peers with a middle SES background in numerical estimation proficiency. In four studies, the at-risk children did not reach the performance level of their age peers, although they improved or had even better gains in their learning than TYP children did (Fuchs, Fuchs, & Karns, 2001; Fuchs et al. 2005; Räsänen et al., 2009; Sood & Jitendra, 2011). Six studies applied delayed measurement. The intervention effect held in five studies, with measurement conducted three to nine weeks after the intervention, and faded in one study, with measurement conducted six months after the intervention.

Discussion

This review focused on providing evidence of the effectiveness of early numeracy interventions for young children at risk for MD, and identified the pedagogical components of the interventions likely to influence the effectiveness. The review yielded 19 intervention studies (five core instruction and 14 supplemental instruction), which provided evidence that early numeracy interventions can effectively improve the numeracy skills of young children at risk for MD. In the majority of the studies, the children receiving interventions outperformed the children in control groups, with the magnitude of effect sizes varying from small to large. Therefore, rather than waiting to provide effective mathematics interventions at school (e.g., Baker et al., 2002; Slavin & Lake, 2008), evidence-based programmes could be used to promote early numeracy skills for low-performing children before the onset of school. If the majority of children could master key early numeracy skills at the beginning of formal schooling, better mathematics learning outcomes should result later on (e.g., Aubrey et al., 2006; Aunola et al., 2004; Aunio & Niemivirta, 2010), and reduce the need for supplemental mathematics support at school age. However, research suggests that there will always be a lack of universal response, meaning that there will be a portion of children who will not respond to intervention, regardless of the evidence-based interventions used (Fuchs, Fuchs, & Compton, 2012).

Effectiveness of Early Numeracy Interventions

A variety of research designs guided the primary studies. Hence, interpreting and comparing the intervention effects between studies was not straightforward. In addition to studies applying only pre-post designs, some studies also used delayed post-tests. Control groups included active (CA) and passive controls (CP), as well as groups receiving other types of mathematics interventions (CI). Comparisons between groups receiving intervention and passive controls have previously shown greater effects than studies having active controls. In the current review, however, the mean magnitudes of effects in both comparisons were of medium size (intervention vs. CA mean g = 0.76 and intervention vs. CP mean g = 0.62). Comparisons to passive control groups do not differentiate between intervention effects and the possible effects simply from the extra attention given to children. Therefore, including an active group, in addition to passive controls, is recommended. On the current review, only one study used both active and passive controls (Jordan et al., 2012). In that study, the experimental group outperformed both control groups, suggesting that the extra attention given to children was not the reason for the group differences. The current review required a passive or active control group in studies using several mathematics interventions, as the absence of such a group makes it difficult to determine if the improvements are due to normal development or to the intervention (Jordan et al., 2012). For example, comparing the Building Blocks intervention group (Clements & Sarama, 2008) to CA (g = 1.11), produced a larger effect than a comparison to CI (g = .56), suggesting that improvements were due to Building Blocks intervention. Furthermore, in the CAI study (Räsänen et al., 2009), the Number Race (NR) and Graphogame Math (GM) were equally effective at improving children's comparison skills, when comparing only these groups. However, the children playing GM made statistically significant, bigger gains in their comparison skills, compared to TYP children, than children playing the NR. As another example, the Great Race board game enhanced children's numerical estimation proficiency if the numbers on the board were organized linearly, not in a circular form. The experimental children outperformed both CI and CA groups in number line estimation skills; in addition, the experimental group outperformed CA in numerical magnitude comparison. In summary, it would be important to include an active control group, with children at a similar performance level, in the intervention design rather than using only a passive control group.

None of the core intervention studies included delayed post measurements. Hence, we have no evidence to support whether the positive intervention effects held. In supplemental interventions, six studies applied a delayed post-test measurement after the intervention had finished. In five short intervention studies, the intervention effect held when measured between three to nine weeks after the intervention. For example, with two months of Number Sense Intervention (Dyson et al., 2011; Jordan et al., 2012), the improved skills in number recognition, number combinations, calculation, and story problems were maintained 6–8 weeks later. In one long intervention (Aunio et al., 2005), the effect faded after six months from the end of the intervention. The longer the positive effect of early numeracy intervention remains after the intervention, the more effective it is likely to be at preventing mathematics difficulties. Hence, it is important to include delayed post-tests in intervention studies.

As one of the aims of interventions is to reduce or even close the performance gap (Fuchs, 2011), a group of TYP children, or standardized tests, should provide benchmarks for typical development, and should be included in intervention studies to show the performance level of LOW children compared to TYP children after the intervention. Although most of the core interventions boosted at-risk children's numeracy skills significantly and reduced the gap to age-related peers, only one core intervention study (Clarke et al., 2011) reported the percentage of at-risk children who reached age-level performance. Almost half of the supplemental intervention studies provided benchmarks if the at-risk children were able to reduce or close the performance gap between themselves and typically performing children. Despite at-risk children were able to reduce the performance gap by making remarkable progress, several studies reported that at-risk children still lagged behind the performance level of their age related peers. One study (Bryant et al., 2011) reported that almost half of the at-risk children moved out of the risk category, compared to one fifth of the controls. Two studies (Aunio et al., 2005; Siegler & Ramani, 2008) provided promising signs that interventions can even close performance gaps. Due to the small sample sizes of these two studies, the results are only suggestive.

An advantage of using evidence-based interventions, compared to non-research-based instruction, is that, with frequent progress monitoring, they can guide how at-risk children's mathematics learning should be supported. When applying evidence-based interventions and when using them to guide the level of support, it is important to consider the types of learners for whom the intervention was originally designed to provide support (Klinger & Edwards, 2006), and to implement it with fidelity (Lembke, Hampton, & Byers, 2012). If a child does not respond to the core instruction as expected, on the basis of the research evidence, more intensive supplemental instruction should be provided. If the supplemental intervention helps the child to close the performance gap, then the child could move back to receiving only core instruction with regular follow-ups to ensure that the child is keeping up with the learning pace. If, on the other hand, the child makes good progress receiving supplemental instruction, but does not close the performance gap, it would be important to carry on the supplemental instruction with possible intervention modifications (e.g., duration, content).

Influence of Pedagogical Components

Setting

The positive intervention effects indicate that four- to five-year-old at-risk children's numeracy skills can be successfully promoted in whole group settings, consistent with the suggested viewpoint of Clarke, Doabler, et al. (2011). This has many benefits. If at-risk children's learning can be effectively promoted together within regular instruction instead of pullout lessons, only one teacher is needed to provide instruction. From the children's point of view, time for the intervention is not taken out of time allocated for other subjects, as can be the case in supplemental instruction. Although the instruction was given in a whole group setting, it did not mean that children were required to work as a whole group entity all the time. Children had also opportunities to work in small groups and in pairs, or individually with computers. Although the amount of differentiation used in the classrooms was not adequately reported in the primary studies, the arrangements within the classroom, including small group work and CAI, should provide opportunities for differentiation (Gersten, Beckmann, et al. 2009), allowing teachers to provide appropriate guidance and activities for children, based on their performance levels.

Small-group instruction was the primary group setting in half of the studies assessing supplemental instruction. Working with a small group of children or one-to-one, a teacher has the opportunity to pay more attention to individual children's needs and to guide, model, and give personal feedback. Individual one-to-one instruction was used with interventions involving CAI or while playing board games. One concern related to supplemental interventions provided in small groups or individually is that teachers must be given the time and resources for implementation (Fuchs et al., 2012).

Duration

Core instruction interventions replaced the typical mathematics instruction used in the classroom during the school year. Teachers engaged in following the programme as intended, mostly with daily sessions. When implementing such a long intervention, there might be a temptation to finish or to modify the programme, if children do not show improvements in a short time, or if the teaching approach is different from the one familiar to the teacher. In supplemental interventions, the duration varied greatly. However, even short interventions of less than 12 weeks, implemented within regular mathematics instruction time, or as short pullout sessions using either CAI or board games, significantly improved children's numeracy skills. As brief interventions are low in external validity (Slavin, 2008), caution must be used in generalizing the results.

Numeracy content used for intervention training and progress measurement

A wide range of numeracy skills was covered in the interventions, despite the relatively few intervention studies found. Interventions included numeracy skills such as verbal and object counting, subitising, addition and subtraction skills, comparison skills, and numeracy-related logical abilities. Studies used either standardized or developer-made measures, or both. As developer-made measures tend to measure the outcomes taught in the intervention more specifically and reveal more positive effect sizes compared to using standardized measures (Slavin, 2008), caution should be applied when comparing the studies based on effect sizes. In core instruction, greater effect sizes were evident when using only developer-made tests (BB interventions) compared to studies that used only standardized tests or both standardized and developer-made tests (ELM intervention). Skill-specific, supplemental intervention studies applied only developer-made tests, which may partly explain the large effect sizes obtained.

Conductor and professional developmental support

Teachers implemented core instruction, whereas either teachers or members of the research team implemented the supplemental instruction intervention. Using teachers as conductors improves the ecological validity of the intervention as it is conducted in real school life. However, teachers who volunteer to participate as conductors of the interventions may be more motivated compared to those in control groups, possibly affecting the outcomes (Slavin, 2008). In the primary studies, the members of the research team who conducted the interventions were additional resource personnel for preschools and kindergartens. One might question whether the interventions conducted by members of the research team would achieve similar magnitudes of effects if conducted by teachers, and whether it would even be possible to conduct the interventions with the schools' current resources. Short supplemental interventions or using computers as intervention tools might be transferred to in-school intervention implementation with relative ease, whereas longer interventions often require additional personnel and resources. Professional developmental training was emphasized in core instruction and was provided in small amounts for the conductors of the supplemental interventions. Although teachers' knowledge about mathematics learning and teaching is essential for effective implementation of interventions (Haseler, 2008), the amount of tutoring required for implementing core instruction or long supplemental interventions can be a feasibility obstacle for some teachers and schools, especially if tutoring is chargeable or if tutoring is provided outside the school.

Instructional design features

Progress in mathematics learning was evident when instruction included one or more of the following instructional features: explicit instruction, peer-assisted instruction, CRA, CAI, or games. Previously, Malofeeva's (2005) findings supported using a combination of guided and explicit approach in teaching young children mathematics. Interventions in the current review applied mostly explicit approaches, which have been found to be effective with low-performing school-aged children (e.g., Kroesbergen & Van Luit, 2003). The findings on peer-assisted instruction were limited to one study. Malofeeva (2005) found peer-assisted tutoring working well with young children. The findings regarding the benefits of peer-assisted tutoring with low-performing school-aged children have been incoherent (Baker et al., 2002; Kroesbergen & Van Luit, 2003).

Findings concerning the effectiveness of the CAI have also been contradictory, either less effective than instruction given by teachers (Kroesbergen & Van Luit, 2003, Malofeeva, 2005), or very effective (Xin & Jitendra, 1999). About 30% of the supplemental interventions applied the CRA approach. Hands-on materials often used with young children in teaching mathematics should provide good opportunities to begin practising from the concrete level, and then progress to pictorial and symbolic representations. Games provided either one part of the instruction or were the only means of providing positive learning outcomes.

Limitations

To increase the validity of the review, we followed strict criteria in selecting studies. The criteria were, however, not as tight as suggested by Slavin (2008). Including only research published in English might have overlooked some relevant studies. The review also excluded the results of intervention studies in dissertations and evaluations published in non-peer-reviewed journals; including only published studies may have increased the likelihood of the file-drawer effect, as publication policies tend to favour studies reporting statistically significant results (Ellis, 2010).

Future research

As relatively few studies were found for this review, more intervention studies focusing on enhancing young children's numeracy skills are required in the future. To meet the standards of high-quality intervention studies, researchers should place more emphasis on conducting randomized large-scale studies with different types of control groups, with adequate duration, and preferably, with standardized measures. Delayed follow-up measurements should be applied, as they are important in providing evidence about longer-lasting intervention effects. The longer the effect remains, the more effective the intervention can be at preventing mathematics difficulties. No interventions were found for seven-year-olds, and only four were available for 6-year-olds, although the age is crucial in mathematics learning. More intervention research to meet the needs of this age group is needed in the future to provide teachers with evidence-based practices to promote the learning of low-performing children successfully.

References

References marked with an asterisk indicate studies included in the review analysis.

Anthony, G. & Walshaw, M. (2009). *Effective pedagogy in mathematics*. Educational practices series -19. Belgium: International Academy of Education. Retrieved from http://www.ibe.unesco.org/fileadmin/user_upload/ Publications/Educational_Practices/EdPractices_19.pdf.

*Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math development in head start classrooms. *Journal of Educational Psychology, 94*(4), 762–770. doi:10.1037//0022-0663.94.4.762

Aubrey, C., Dahl, S., & Godfrey, R. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal, 18*(1), 27–46.

Aunio, P., Hautamäki, J., Sajaniemi, N., & Van Luit, J. E. H. (2009). Early numeracy in low-performing young children. *British Educational Research Journal, 35*(1), 25–46. doi:10.1080/01411920802041822

*Aunio, P., Hautamäki, J., & Van Luit, J. E. H. (2005). Mathematical thinking intervention programmes for preschool children with normal and low number sense. *European Journal of Special Needs Education, 22*(2), 131–146. doi:10.1080/08856250500055578

Aunio, P. & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20*(5), 427–435.

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.

Baker, S., Gersten, R., and Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal, 103*, 51–73.

*Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Hughes Pfannelstiel, K., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children, 78*(1), 7–23.

Bryant, D. P., Roberts, G., Bryant, B. R., & DiAndreth-Elkins, L. (2011). Tier 2 early numeracy number sense interventions for kindergarten and first-grade students with mathematics difficulties. In R. Gersten & R. Newman-Gonchar (Eds.), *Understanding RTI in mathematics. Proven methods and applications* (pp. 65–83). Baltimore, ML: Brookes.

*Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disability Quarterly, 31*, 11–20.

Clarke, B., Doabler, C. T., Baker, S. K., Fien, H., Jungjohann, K., & Strand Cary, M. (2011). Pursuing instructional coherence. In R. Gersten & R. Newman-Gonchar (Eds.), *Understanding RTI in mathematics*. *Proven methods and applications* (p. 49–64). Baltimore, ML: Brookes.

*Clarke, B., Smolkowski, K., Baker, S., Fien, H., Doabler, C. T., & Chard, D. J. (2011). The impact of a comprehensive Tier 1 core kindergarten program on the achievement of students at risk in mathematics. *The Elementary School Journal, 111*(4), 561–584.

*Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*(2), 136–163.

*Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a researched-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443–494. doi:10.3102/0002831207312908

*Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. *Journal for Research in Mathematics Education, 42*(2), 127-166.

Codding, R. S., Hilt-Panahon, A., Panahon, C. J., & Benson, J. L. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and Treatment of Children, 32*(2), 279–312.

Cohen, J. (1988). *Statistical power analysis for the behavioral science*s (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Desoete, A., Ceulemans, A., DeWeerdt, F., & Pieters, S. (2010). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Psychology, 82*, 64–81.

Desoete, A., Stock, P., Schepens, A., Baeyens, D., & Roeyers, H. (2009). Classification, seriation, and counting in grades 1, 2, and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *Journal of Psychoeducational Assessment, 27*(3), 252–264. doi:10.1177/0734282908330588

*Dyson, N. I., Jordan, N. C., & Glutting, J. (2011). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. *Journal of Learning Disabilities*. Published online. doi: 10 1177/0022219411410233.

Ellis, P. D. (2011). *The essential guide to effect sizes. Statistical power, meta-analysis, and the interpretation of research results.* Cambridge, UK: Cambridge University press.

Fuchs, L. S. (2011). *Mathematics intervention at the secondary prevention level of a multi- tier prevention system: Six key principles*. Retrieved from: http://www.rtinetwork.org/essential/tieredinstruction/tier2/mathintervention.

*Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*(3), 495–513.

Fuchs, D. & Fuchs, L. S. (2006). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children, 38*(1), 57–61.

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., & Karns, K. (2001). Enhancing kindergartners' mathematical development: Effects of peer-assisted learning strategies. *The Elementary School Journal, 101*(5), 495–510.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
Retrieved from http://ies. ed.gov/ncee/wwc/publications/practiceguides/.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (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.

Haseler, M. (2008). Making intervention in numeracy more effective in schools. In A. Dowker (Ed.), *Mathematical Difficulties* (pp. 225–241). San Diego, CA: Elsevier.

*Jordan, N. C., Glutting, J., Dyson, N., Hassinger-Das, B., & Irwin, C. (2012). Building kindergartners' number sense: A randomized controlled study. *Journal of Educational Psychology, 104*(3), 647–660. doi: 10.1037/a0029018.

Jordan, N.C., Kaplan, D., Olah, L., & Locuniak, M. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*, 153–175.

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

*Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness, 1*, 155–178. doi:10.1080/19345740802114533.

Klinger, J. K., & Edwards, P. A. (2006). Cultural considerations with response to intervention models. *Reading Research Quarterly, 1*, 108–117. doi:10.1598/RRQ.41.1.6

Krajewski, K. & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction, 19*, 513–526. doi:10.1016/j.learninstruc.2008.10.002.

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

Kunsch, C. A., Jitendra, A. K., & Sood, S. (2007). The effects of peer-mediated instruction in mathematics for students with learning problems: A research synthesis. *Learning Disabilities Research and Practice, 22*(1), 1–12.

Lembke, E. S., Hampton, D., & Beyers, S. J. (2012). Response to intervention in mathematics: Critical elements. *Psychology in the Schools, 49*(3), 257-272. doi:10.1002/pits.21596

Malofeeva, E. V. (2005). *Meta-analysis of mathematics instruction with young children* (Unpublished doctoral dissertation).University of Notre Dame, Notre Dame, IN.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*(4), 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x.

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*, 306–321.

Morris, S. B. (2008). Estimating effect sizes from pre-test-post-test-control group designs. *Organizational Research Methods, 11*(2), 364–386. doi:10.1177/1094428106291059.

O'Connor, D., Green, S., & Higgins, J. P. T. (2008). Defining the review question and developing criteria for including studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 83–94). Chichester, England: Wiley-Blackwell.

Price, G. R. & Ansari, D. (2013). Dyscalculia: Characteristics, causes, and treatments. *Numeracy, 6*(1), 1-16.

*Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394. doi:10.1111/j.1467-8624.2007.01131.x.

Riccomini, P. J., & Smith, G. W. (2011). Introduction of response to intervention in mathematics. In R. Gersten & R. Newman-Gonchar (Eds.), *Understanding RTI in Mathematics*. *Proven methods and applications* (pp. 1–16). Baltimore, ML: Brookes.

*Räsänen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development, 24,* 450–472. doi:10.1016/j.cogdev.2009.09.003.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research*. *Learning trajectories for young children*. New York, NY: Routledge.

*Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11*(5), *655*–661. doi:10.1111/j.1467-7687.2008.00714.x.

*Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games - but not circular ones - improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560. doi:10.1037/a0014239.

Slavin, R. E. (2008). What works? Issues in synthesizing educational program evaluations. *Educational Researcher, 37*(1), 5–14. doi:10.3102/0013189X08314117.

*Sood, S. & Jitendra, A. K. (2011). An exploratory study of a number sense program to develop kindergarten students' number proficiency. *Journal of Learning Disabilities*. Published online. doi:10.1177/002219411422380.

Turner, H. M. & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders, 33*, 42–55.

What Works Clearinghouse (2008). *Procedures and Standards Handbook (version 2.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_standards_handbook.pdf.

Witzel, B. S., Mink, D. V., & Riccomini, P. J. (2011). Using visual representations to instruct and intervene with secondary mathematics. In R. Gersten & R. Newman-Gonchar (Eds.), *Understanding RTI in Mathematics*. *Proven methods and applications* (pp. 151–167). Baltimore, ML: Brookes.

Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical Word problems for students with learning problems: A meta-analysis. *Journal of Special Education, 32*(4), 207–225.

Appendix

Appendix A

Overview and Effects of Intervention Studies Included in the Review

Table A1

Descriptive Statistics and Effect Sizes of Core Instruction Intervention Studies

	Sample				Intervention							
Study	Age	At-risk status	Participants (N)	Design	Programme/content	Duration	Group setting	Leader	Instructional design feature	Measure	Pre-post Effect size (*g*) with 95% CI	Follow-up (YES/NO)
Chard, Baker, Clarke, Jungjohann, Davis, & Smolkowski (2008)	5	LOW	E: 186 (TYP&LOW) CA: 102 (TYP&LOW)	QE	E: Early Learning in Mathematics (ELM) CA: typical math activities	25 weeks (3000 mins)	W	T	explicit and guided instruction, pair activities, CRA	SESAT-2 (all) [a)]	0.33 [0.09,0.57]	NO
Clarke, Smolkowski, Baker, Fien, Doabler, & Chard (2011)	5	LOW	E: 660 (incl. 313 LOW) CA: 553 (incl. 342 LOW)	RA	E: Early Learning in Mathematics (ELM) CA: typical math activities	school year (7200 mins)	W	T	explicit and guided instruction, pair activities, CRA	TEMA (all) EN-CBM (all) TEMA (TYP) EN-CBM (TYP) TEMA (LOW) EN-CBM (LOW)	0.24 [0.13,0.35] 0.24 [0.13,0.36] 0.03 [-0.18,0.23] 0.04 [-0.17,0.24] 0.36** [0.21,0.51] 0.36* [0.21,0.51]	NO
Clements & Sarama (2007)	4	SES	E: 30 CA: 38	QE	E: Building Blocks CA: typical math activities	25 weeks (daily sessions)	W	T	explicit instruction, small-group instruction and CAI	Building Blocks Assessment: Number Geometry	 1.11* [0.60,1.62] 2.45** [1.82,3.08]	NO
Clements & Sarama (2008)	4	SES	E: 101 CA: 101 CI: 51	RA	E: Building Blocks CA: typical math activities CI: Preschool Mathematics Curriculum	26 weeks (2470 mins)	W	T	explicit instruction, small-group instruction and CAI	Early Mathematics Assessment (EMA): Number & Geometry E vs. CA E vs. CO	 1.11*** [0.82,1.41] 0.56** [0.21,0.89]	NO

Clements, Sarama, Spitler, Lange, & Wolfe (2011)	4	SES	E: 927 CA: 378	RA	E: Building Blocks CA: typical math activities (different curricula)	school year	W	T	explicit instruction, small-group instruction and CAI	The Research-Based Elementary Math Assessment (REMA): Number & Geometry	 0.77*** [0.59,0.84]	NO

Note. At risk status: L = low-performing, S = low socio-economic status; T = typically performing; Participants: E = experimental group, CA = active control group, CI = other math intervention; Design: RA = random assignment, QE = quasi-experimental; Group setting: W = whole group; Leader: T = teacher; CRA = concrete-representational-abstract model, CAI = computer-assisted instruction; Effect size: calculated as unbiased Hedges' *g* with 95% confidence intervals, posttest scores adjusted for pretest scores. [a)] intervention did not differ for LOW and TYP children ($p = .599$).

* $p < .05$, ** $p < .01$, *** $p < .001$, based on *p*-values reported in the primary study

Table A2

Descriptive Statistics and Effect Sizes of Supplemental Instruction Intervention Studies

	Sample				Intervention							
Study	Age	At-risk status	Participants (N)	Design	Programme/content	Duration	Group setting	Leader	Instructional design feature	Measure	Pre-post Effect size (*g*) with 95% CI	Follow-up (YES/NO)
Arnold, Fisher, Doctoroff, & Dobbs (2002)	4	SES	E: 56 CA: 56	RA	E: Emergent math skills CA: typical math activities	6 weeks (daily sessions), replacing typical activities	W	T	math incorporated into daily routines (e.g., circle time and small-group instruction)	TEMA-2	0.45** [0.08,0.83]	NO
Aunio, Hautamäki, & Van Luit (2005)	5	LOW	E: 22 (incl. 5 LOW) CP: 23 (incl. 7 LOW)	RA	Let's Think! and Young Children with special educational needs count, too!	36 weeks (1500 mins)	S	T, R	explicit instruction	Early Numeracy Test: Relational (all) Counting (all) Relational (LOW) Counting (LOW)	 0.58* [-0.01,1.16] 0.87* [0.27,1.47] *ns* *, not fully reported	YES, effects faded after 6 months
Bryant, Bryant, Roberts, Vaughn, Hughes Pfannenstiel, Porterfield, &	6	LOW	E: 139 CP: 64	RA	Early numeracy Tier 2 intervention	19 weeks (1900 mins)	S	R	explicit and guided instruction, CRA	TEMI-PM: Magnitude comparison Place value +/- Combinations Number sequences Total score	 0.21 [-0.09,0.51] 0.44** [0.14,0.74] 0.59** [0.29,0.83] 0.53** [0.23,0.83] 0.55** [0.25,0.85]	NO

Gersten (2011)										TEMI-O: Problem solving Computation Total score SAT-10: Problem solving Procedures Total score	 0 [-0.30,0.29] 0.49** [0.20,0.79] 0.28 [-0.02,0.57] 0.28 [-0.02,0.57] 0.14 [-0.16,0.43] 0.20 [-0.09,0.50]	
Dyson, Jordan, & Glutting (2011)	5	SES	E: 56 CP: 65	RA	Number Sense Intervention	8 weeks (720 mins)	S	R	explicit instruction, pair activities, CRA, games	Number Sense Brief: Counting skills Counting principles Number recognition Number knowledge Nonverbal calculation Story problems Number combinations Total score Woodcock-Johnson: Applied problems Calculation Total score	 0.18 [-0.18,0.54] -0.40 [-0.76,-0.04] 0.73** [0.37,1.10] 0.57** [0.20,0.93] -0.11 [-0.47,0.24] 0.97** [0.59,1.34] 0.56* [0.20,0.92] 0.68** [0.31,1.04] 0.17 [-0.18,0.53] 0.44** [0.08,0.80] 0.32* [-0.04,0.68]	YES, effects held after six weeks
Fuchs, Fuchs, & Karns (2001)	5	LOW	E: 79 (incl. TYP high: 14, TYP med: 49, LOW: 8, SEN: 8) CA: 83 (incl. TYP high: 17, TYP med: 52, LOW: 7, SEN: 7)	RA	E: PALS: Peer assisted learning strategies CA: typical math activities	15 weeks (600 mins), replacing part of the typical classroom instruction	W	T	peer assisted activities	SESAT: all TYP high [a)] TYP med LOW SEN	 *, not fully reported -0.43 [-1.13,0.27] 0.54 [0.14,0.93] 0.62 [-0.36,1.60] 0.38 [-0.59,1.34]	NO
Fuchs, Compton, Fuchs, Paulsen, Bryant, & Hamlett (2005)	6	LOW	E: 64 CP: 63, in addition C2: 437 TYP	RA	Preventive first grade tutoring in mathematics and Math Flash software	16 weeks (1920 mins)	S	R	explicit instruction, CRA, CAI	CBM computation Basic facts + Basic facts - Woodcock-Johnson: Calculation Applied problems Grade 1 concepts Story problems	1.16* [0.79,1.54] 1.42 [1.04,1.81] 0.61 [0.26,0.97] 0.50* [0.15,0.85] 0.10 [-0.25,0.44] 0.92* [0.56,1.29] 1.08 [0.71,1.45]	NO

										TYP outperformed LOW in all posttest measures	
Fuchs, Fuchs, Hamlet, Powell, Capizzi, & Seethaler (2006)	6	LOW	E: 16 CA: 17	RA	E: Math Flash software CA: Spelling software	18 weeks (500 mins)	S	R	CAI (children working individually with computers)	Addition fact fluency Subtraction fact fluency Story problems (transfer task)	1.01* [0.30,1.72] -0.01 [-0.68,0.66] -0.13 [-0.79,0.54]	NO
Jordan, Glutting, Dyson, Hassinger-Das, & Irwin (2012)	5	SES	E: 44 CA: 44 CP: 44	RA	E: Number Sense Intervention CA: Language intervention	8 weeks (720 mins)	S	R	explicit instruction, pair activities, games	E vs. CA: [b)] Number Sense Brief: Counting skills Number recognition Number comparisons Nonverbal calculation Story problems Number combinations Total score Woodcock-Johnson: Applied problems Calculation Total score E vs. CP: Number Sense Brief: Counting skills Number recognition Number comparisons Nonverbal calculation Story problems Number combinations Total score Woodcock-Johnson: Applied problems Calculation Total score	 0.18 [-0.24,0,59] 0.87 [0.43,1.30] 0.53 [0.10,0.95] -0.02 [-0.44,0.39] 2.04 [1.53,2.56] 1.50 [1.03,1.97] 1.01 [0.57,1.45] 0.25 [-0.16,0.67] 2.53 [1.97,3.09] 1.11 [0.66,1.55] 0 [-0.41,0.41] 0.91* [0.48,1.35] 0.79* [0.36,1.22] 0.22 [-0.20,0.63] 2.27*** [1.73,2.80] 1.40*** [0.94,1.87] 1.16** [0.71,1.61] 0.23 [-0.18,0.65] 2.24*** [1.71,2.77] 1.09** [0.64,1.53]	YES, effects held after 8 weeks

Klein, Starkey, Clements, Sarama, & Iyer (2008)	4	SES	E: 139 CP: 139	RA	Pre-K Mathematics and DLM Express math software	29 weeks (1160 mins)	S	T	explicit teaching, CAI, home activities	Child Math Assessment	***0.50 [0.26,0.74]	NO
Ramani & Siegler (2008) (experiment1)	4	SES	E: 68 CA: 56	RA	The Great Race board game, E: with numbers, CA: with colours	2 weeks (80 mins)	I	R	games	Numerical identification Numerical magnitude comparison Counting Number line estimation: Linearity Slope Accuracy	0.66*** [0.30,1.02] 0.96*** [0.59,1.33] 0.67*** [0.31,1.03] 1.09*** [0.71,1.47] 1.03*** [0.66,1.40] not fully reported	YES, effects held after 9 weeks
Räsänen, Salminen, Wilson, Aunio, & Dehaene (2009)	6	LOW	E: 15 CI: 15, in addition CP: 30 TYP	RA	E: Number Race software CI: Graphogame-Math software	3 weeks (225 mins)	I	T	CAI	Verbal counting Number comparison Subitising Object counting Arithmetic	-0.32 [-1.04,0.40] -0.10 [-0.78,0.60] 0.27 [-0.43,0.97] -0.24 [-0.93,0.46] 0.38 [-0.32,1.08] CI improved sig. more than TYP in number comparison; TYP outperformed E and CI in all measures	YES, after 3 weeks no differences found between E vs. CI; CI had improved sig. more than TYP in number comparison
Siegler & Ramani (2008)	4	SES	E: 18 CA: 18 in addition CP: 22 TYP	RA	The Great Race board game E: with numbers, CA: with colours	2 weeks (60 mins)	I	R	games	Number line estimation: Linearity Slope Accuracy Numerical magnitudes	1.58*** [0.85,2.32] 1.30*** [0.59,2.01] not fully reported not fully reported E vs. TYP (middle-SES): no difference in posttest results	NO

Siegler & Ramani (2009)	4	SES	E: 30 CI: 29 CA: 29	RA	The Great Race board game E: linear board with numbers, CI: circular board with numbers, CA: numerical activities	3 weeks (100 mins)	I	R	games	E vs. CI: Numerical identification Numerical magnitude comparisons Counting Number line estimation: Linearity Slope Accuracy E vs. CA: Numerical identification Numerical magnitude comparison Counting Number line estimation: Linearity Slope Accuracy	 ns ns ns 0.64* [0.13,1.16] 0.70** [0.18,1.22] 0.67* [0.15,1.18] ns 0.74** [0.22,1.26] ns 0.75** [0.22,1.27] 0.87** [0.34,1.39] 0.62* [0.10,1.14]	NO
Sood & Jitendra (2011)	5	SES	E: 61 CA: 40	QE	Number Sense Instruction (NSI)	4 weeks (400 mins), replacing part of the typical classroom instruction	W	T	explicit instruction in whole and small-groups, pair activities	EN-CBM: Oral counting fluency Counting from Number identification Number Sense: spatial relationships More/less relationships Benchmarks (5 and 10) Nonverbal calculations Total score [c)]	 0.10 [-0.30,0.50] 0.31 [-0.09,0.71] 0.32* [-0.08,0.72] 0.55*** [0.15,0.95] 1.14** [0.72,1.57] 0.87*** [0.45,1.28] 0.68* [0.27,1.08] 1.08 [0.66,1.51] The effects for NSI were not mediated by at-risk status for MD[d)]	YES, effects held after 3 weeks

Note. At risk status: LOW = low-performing, SES = low socio-economic status; SEN = special educational needs; T = typically performing; Participants: E = experimental group, CA = active control group, CP = passive control group, CI = other math intervention control group; Design: RA = random assignment, QE = quasi-experimental; Group setting: W = whole group, S = small group, I = individual; Leader: T = teacher, R = member of the research team; CRA = concrete-representational-abstract model, CAI = computer-assisted instruction; Effect size: calculated as unbiased Hedges' *g* with 95% confidence intervals, posttest scores adjusted for pretest scores. [a)] *p*-values were not reported for subgroups. [b)] *p*-values were not reported for E vs. CA. Differences between CA and CP were *ns*. [c)] *p*-value for total score was not reported. [d)] The sample was divided as children at-risk or not at-risk for MD before the intervention, using cut-off criteria of 40th percentile on SESAT 1 screening measure.

* $p < .05$, ** $p < .01$, *** $p < .001$, based on *p*-values reported in the primary study

APPENDIX B
Excluded Intervention Studies

Authors	Title	Reason not included
Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E (2012)	Fostering at-risk kindergarten children's number sense	No control group (two intervention groups)
Baroody, A. J., Eiland, M., & Thompson, B. (2009)	Fostering at-risk preschoolers' number sense	No control group (several intervention groups), insufficient data
Bermejo, V., Morales, S., & Garcia de Osuna, J. (2004)	Supporting children’s development of cardinality understanding	Different outcome measures from pretest, measure inherent to treatment
Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N., & Chavez, M. M. (2008)	Mathematics intervention for first- and second-grade students with mathematics difficulties. The effects of Tier 2 intervention delivered as booster lessons	Insufficient data
Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N. N., Funk, C., Winter, A., Shih, M., & Pool, C. (2008)	The effects of Tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties	Insufficient data
Carr, M., Taasoobshirazi, G., Stroud, R., & Royer, J. M. (2011)	Combined fluency and cognitive strategies instruction improves mathematics achievement in early elementary school	Sample of typically developing children
Codding, R. S., Chan-Iannetta, L., George, S., Ferreira, K., & Volpe, R. (2011)	Early number skills: Examining the effects of class-wide interventions on kindergarten performance	At-risk cut-off score too high (mean as a cut-off criteria), insufficient data for at-risk children
Coles, C. D., Kable, J. A., Taddeo, E. (2009)	Math performance and behavior problems in children affected by prenatal alcohol exposure: Intervention and follow-up	Sample of special needs children
Greenes, C., Ginsburg, H., & Balfanz, R. (2004)	Big math for little kids	Qualitative description
Griffin, S. (2004)	Building number sense with Number Worlds: a mathematics program for young children	Qualitative description
Kaufmann, L., Delazer, M., Pohl, R., Semenza, C., & Dowker, A. (2005)	Effects of a specific numeracy educational program in kindergarten children: A pilot study	Sample of typically developing children
Lai, M.-L., Baroody, A. J., & Johnson A. R. (2008)	Fostering Taiwanese preschoolers’ understanding of the addition-subtraction inverse principle	No controlling for pretest, measure inherent to treatment
Opel, A., Zaman, S. S., Khanom, F., & Aboud, F. E. (2012)	Evaluation of a mathematics program for preprimary children in rural Bangladesh	No pretest data, or posttest data for cumulative final test
Ramani, G. B. & Siegler, R. S. (2011)	Reducing the gap in numerical knowledge between low- and middle-income preschoolers	Insufficient data
Starkey, P., Klein, A., & Wakeley, A. (2004)	Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention	No pretest data for controls
Van Luit, J. E. H. & Schopman, E. A. M. (2000)	Improving early numeracy of young children with special educational needs	Sample of special needs children
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	A two-period cross-over design (no actual control group)
Young-Loveridge, J. M. (2004)	Effects on early numeracy of a program using number books and games	Insufficient data