Ulster University

In collaboration with

Loughborough University Mathematics Education Centre

UCD Dublin

Interventions to improve mathematical achievement in primary school-aged children

A Systematic Review

Simms, V., McKeaveney, C., Sloan, S. & Gilmore, C.

ulster.ac.uk

Nuffield Foundation

00553-05_UU_A4_Report_v7.indd 1 05/06/2019 15:20

About the authors

Victoria Simms is a Reader in Developmental Psychology and Research Director of the School of Psychology at Ulster University, UK. Her research focuses on the development of mathematical understanding in children and adults. Dr Simms is a chartered psychologist with the British Psychological Society.

Clare McKeaveney is a Post-Doctoral Research Fellow in the School of Medicine at Queen's University Belfast, UK. Her research focuses on the application of rigorous methods in understanding psychological outcomes following adverse health experiences.

Seaneen Sloan is an Assistant Professor in Education in the School of Education at University College Dublin, Ireland. Her research focuses on social and emotional wellbeing and well-being within the school context.

Camilla Gilmore is a Reader in Mathematical Cognition at the Mathematics Education Centre, Loughborough University, UK. Her research focuses on the development of mathematical skills and their impact on achievement. Dr Gilmore is a chartered psychologist and an Associate Fellow of the British Psychological Society.

Table of Contents

00553-05_UU_A4_Report_v7.indd 3 05/06/2019 15:20

Executive Summary

Mathematical achievement is important for children's future educational success, employment opportunities and health outcomes. However, it is recognised that there is substantial underachievement in this subject (e.g. DE, 2015). There is a growing body of literature that assesses the impact of interventions on mathematical achievement (DfE, 2012)

This review of published studies rigorously investigated the outcomes of classroom-based mathematical interventions that were targeted at primary school-aged children. The review assessed whether the interventions had an effect on mathematical achievement (as measured by variety of outcome measures) or related attitudinal or affective constructs (such as mathematical anxiety).

The review had a clearly defined method to identify studies. The criteria for inclusion and exclusion for the study were also defined before the review began. This plan was recorded on the internet prior to beginning the review process so that readers can be assured that the work is rigorous.

Some definitions

What do we mean by interventions?

In this review the term 'intervention' is defined as a deviation from existing teaching practice. Interventions aimed at parents or requiring parental involvement, or those implemented in afterschool or home settings, are beyond the scope of this review and are not included.

How are mathematical disabilities defined in this review?

This review focuses on studies that were not targeted towards children who have been identified as having specific difficulties in mathematics. There are a wide variety of criteria that may be applied to identify a child as having a mathematical disability. In this review if a study screened children against a set criteria (e.g. mathematics achievement below the 25th percentile on standardized mathematical tests) in order to be included in the study the paper was not included in our final summary.

What types of studies were included in this review?

This review included studies that assessed the outcomes of interventions aimed at improving mathematical achievement in primary school-aged children. Forty-five randomized control trials were included along with thirty-five quasi-experimental studies. The studies were published between 2000 and 2017. These studies were mostly conducted in the USA and Europe. Although the randomized control trial studies used random assignment to groups, many did not follow rigorous methods to reduce potential bias.

The review aimed to answer three questions:

1. What types of classroom-based interventions that have been assessed through studies are used with primary school-aged children who do not meet the criteria for mathematics disability?

A wide variety of studies were identified that focused on different topic areas in mathematics, specifically conceptual understanding, understanding magnitudes, knowledge of the number system, fluency with mathematics facts and strategy use. The studies also used a number of different types of teaching methods, such as the use of manipulatives, providing different types of feedback, using computerized environments for learning and varying how information is delivered to pupils (e.g. through song or story books).

2. What are the most effective classroom-based interventions for improving mathematical learning in primary school-aged children who do not meet the criteria for mathematics disability?

Due to the varied nature and targets of the interventions that the review identified we could not compare the studies to make an overall judgement. Directly comparing the effect size of interventions with such varied types, lengths and target areas of mathematics would not lead to valid conclusions. Instead we have summarised the approaches and findings of these studies highlighting the interventions for which rigorous evidence of effectiveness does or does not exist.

3. What are the key characteristics of the most effective interventions for improving mathematical learning in primary-school aged children who do not meet the criteria for mathematics disability (e.g. equipment involved, any associated costs)?

Seventy-two out of the eighty included studies reported a positive impact of the intervention on at least one academic outcome measure. In addition, twelve studies also reported benefits for secondary outcomes, such as attitudes or anxiety related to mathematics. Two studies reported actual costs of interventions. Fourteen studies reported that teachers or assistants would require specialist training to deliver the intervention. Many of the interventions did not require specialist equipment for delivery; most interventions involved routine materials that would be available in most Western-context classrooms. Thirty-six studies required specialist software to deliver the intervention, ten of these studies assessed commercially available software.

Conclusions

The studies identified in this review suggested that there are a number of approaches that have potential to promote mathematical learning in mainstream primary classrooms;

1. Focusing on key topic areas in mathematics such as conceptual understanding, magnitudes and basic number skills.
2. Ensuring that children have a fluent grasp of mathematical facts.
3. Ensuring that children have a wide bank of strategies to complete mathematical problems and that they know when to best apply them.
4. The appropriate use of objects as learning aids.
5. Providing effective and timely feedback.
6. Using technology that has been developed with clear understanding of how children learn.
7. Varying how mathematical content is delivered in the classroom, such as through physical activity or group work.

However, it is important to note that the studies identified in this review were of varying quality. Therefore, we suggest that teachers and policy makers make any decisions for their practice or policy with caution. The lack of replication studies or studies that compare the effectiveness of different interventions is striking. This review highlights the necessity for further, rigorous studies that assess the efficacy and effectiveness of interventions to improve mathematical achievement. The interventions used a wide range of outcome measures (with a high prevalence of author-generated measures) making it difficult to compare across studies.

Recommendations

Policy makers should consider the following recommendations:

1. The development of guidance on interventions should be made with caution as the majority of the evidence that we identified was of low quality.
2. Support teachers through training on evidence based education so that they can evaluate studies that may help inform their practice.
3. Develop a Core Outcome Set (an agreed set of appropriate outcome measures) to aid comparison across studies.

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and qualitative methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit ***www.nuffieldfoundation.org***

Nuffield Foundation

The need for the review

Mathematics skills are increasingly important for everyday life because many activities require some form of mathematical understanding and decision-making (Xing, 2013). Yet, low achievement in mathematics affects a large proportion of the general population. Within the UK it is estimated that 15-17 million adults have poor numeracy skills (Every Child a Chance Trust, 2009; Skills for life, 2012). Research has largely focused on interventions for those with a suspected or diagnosed mathematics learning disability (e.g. dyscalculia). However, there are no consistent standards by which to judge the presence of a mathematics learning disability nor is there agreement concerning the precise definition, criteria, or prevalence (Karagiannakis, Baccaglini-Frank & Papadatos, 2014). A recent report suggests that approximately 5% of the UK population meet the diagnostic criteria for dyscalculia (Department of Work & Pensions, 2017), a smaller scale study in Northern Ireland placed this rate in primary school-aged children at 5.7% (Morysani, van Bers & McCormack, 2018). Individuals with dyscalculia only accounts for a small percentage of the total population (Dowker & Sigley, 2010; Mazzocco, 2015).

A greater number of students struggle with low mathematics performance without a disability diagnosis (Nelson & Powell, 2017). Many children struggle to meet expected levels in mathematics by the end of primary school (DE, 2015). Specifically, Gross (2007) reported that 21% of 11-year-olds leave primary school without reaching the mathematical level expected. Learning mathematics in the primary school years to a large extent determines children's later success in mathematics and as well as other areas such as reading (Sarama & Clements, 2009). Additionally, there are many negative outcomes associated with low numeracy, such as significantly higher risk of unemployment, lower wages, mental health problems, physical illness, arrest and incarceration (Gross, Hudson, and Price, 2009). Moreover, the societal cost of low numeracy is estimated to include billions of pounds of lost revenue in taxes and welfare provision (Every Child a Chance Trust, 2009; OECD 2010). Focusing mathematics education research on children within mainstream classrooms is therefore necessary to provide greater understanding of how to increase mathematical learning across the board.

What impacts on mathematical learning?

Increased investment in mathematics education research has led to a plethora of studies which have identified various skills associated with mathematics performance and proposed a range of potential interventions. Mathematics performance outcomes have been found to be associated with differences in attitudes, motivation (Hammouri, 2004), language ability (Moll, Snowling, Goebel & Hulme, 2015) and IQ (Chu, vanMarle & Geary, 2016), in addition to social and educational factors (Banerjee & Lamb, 2016). Studies have also identified executive functioning skills such as holding and manipulating information in mind (working memory), the ability to suppress distracting information and unwanted responses (inhibition), and the ability to flexibly switch attention between different tasks (shifting) to be associated with mathematics achievement (Cragg & Gilmore, 2014; Cragg, Keeble, Richardson, Roome & Gilmore, 2017).

Additional factors that could cause low achievement in mathematics include anxiety, stress and low self-esteem (Chinn & Ashcroft, 2007). There are multiple factors that contribute to mathematical learning (Wardrop, 2014). Mathematics learning draws on several cognitive mechanisms and memory systems highlighting a highly complex learning process. Therefore, difficulties in mathematics can arise from a wide range of underlying factors. The identification of these skills that underpin mathematical achievement has directly informed the development of experimental interventions to improve mathematical performance that range from one-to-one computerised training to whole-class pedagogical approaches.

Recent research has been dedicated to developing effective interventions to improve mathematical learning in primary school (DfE, 2012). However, this has not necessarily resulted in a clearer understanding of what does and does not improve mathematical learning. Recent reports have called for an increase in systematic and rigorous investigations of education interventions to provide an evidence base to inform decisions about educational changes (DfE, 2013). This type of approach encourages education policy and practice to be guided by the best available evidence.

Evidence-based education: What is it?

Many commentators have called for education policy and practice to be guided by the best available evidence. This is often called "evidence-based education" or "evidence-informed education" (Nelson & Campbell, 2017). The idea is that research studies investigate learning and teaching activities in order to understand learning processes and the factors that impact on children's achievement. This information can then be used to help children learn as effectively as possible. Recently, increased government funding has supported producing this type of research evidence and communicating it broadly to the education community.

00553-05_UU_A4_Report_v7.indd 6 05/06/2019 15:20

Evidence-based education is not, however, straightforward (Gorard, See & Siddiqui, 2017). Some educationalists may be concerned that this approach diminishes the expertise of teachers and the art of teaching or that the findings of research studies are not relevant for the classroom. However, this is not the case. Evidence-based teaching does not generate prescriptive rules (a "recipe book") for teachers to follow (Sharples, 2013). Rather, education research aims to explore and test ideas about how children learn so teachers can make informed decisions about what might work well in their own context.

People often have concerns about how 'generalizable' a research finding is – that is, how far the finding might apply in different situations. It's important to remember that educational research studies are generally conducted to test a theory or an idea about learning or teaching (Mook, 1983). If the study finds evidence to support the theory then it is worth considering how this might apply in different classroom situations. However, it's the theory that should be generalizable, not the individual study findings (Figure 1).

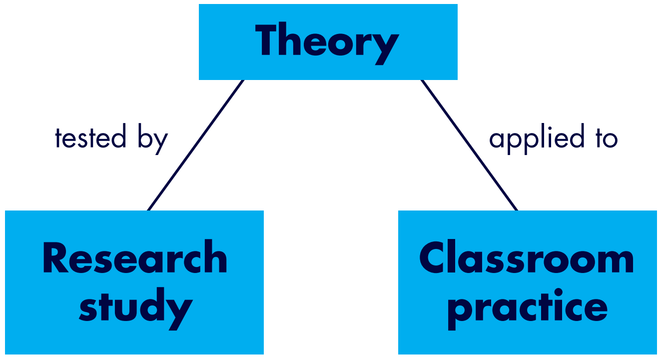


Figure 1: The role of theory in evidence-based education: research studies should be conducted to test a theory of learning and the theory applied to classrooms rather than directly applying the results of a study to the classroom.

The aim of evidence-based education is to provide education professionals with evidence from research about which theories of learning and teaching have received support from research studies. Teachers can then use their expertise and experience to determine how the theory might apply in their classroom.

Different sources of evidence

There are many different types of research methods used to generate information that forms the basis of evidence-based education, including formal or less formal trials of interventions, but also quantitative and qualitative studies of underlying learning theory. These methods all have strengths and weaknesses. However, not all studies are carried out in a rigorous way. Therefore, it is important to take into account research quality; this indicates how much weight the findings should be given. Similar effects that have been found across many studies are likely to be more reliable than results from a single study.

Experimental designs, such as randomised controlled trials (RCTs), are the best and most rigorous way of generating evidence about the effects of interventions (Connolly, Keenan & Urbanska, 2013). RCTs involve randomly allocating individual research participants, or groups of individuals such as classes or schools, to at least two groups. One of these groups receives the intervention, while the other carries on as normal and acts as a 'control group'. At the end of the intervention period, both groups are tested, so comparisons can be made about how much progress each group has made. If the intervention group has made more progress than the control group, this provides evidence that the intervention is having some effect (White, 2013).

Quasi-experimental studies are similar to RCTs, except that they lack the random allocation element (Shadish, Cook & Campbell, 2002). This random allocation is important for helping to ensure that the intervention and control groups are as similar as possible at the beginning of the trial. The more certain researchers can be that the two groups are the same before the intervention increases confidence that any difference between the groups after the intervention is down to the intervention, and not caused by something else (**http://www.consort-statement.org/**). RCTs can also vary in quality so it is important to evaluate their methods and results.

How do we synthesise the evidence base in education?

In order to synthesise evidence researchers carry out a systematic review. A systematic review is a method of collecting together all of the research about a specific question and reviewing it in a way that follows a strict process. Quite often, systematic reviews are conducted to understand the effect of an intervention or programme on certain outcomes, such as achievement, for a specific group of people. Systematic reviews therefore aim to draw conclusions about how effective interventions are based on the best available evidence that is produced by high quality research (Mulrow, 1994).

This is important because a huge amount of research of varying quality is conducted and published, and often, different studies on the same topic or issue may have conflicting findings. This presents a challenge, for researchers, practitioners and the general public, in trying to make sense of what research is telling us about a particular topic, or about certain interventions. This is where systematic reviews of interventions are a useful tool (Davies, 2010).

00553-05_UU_A4_Report_v7.indd 7
05/06/2019 15:20

A systematic review begins by posing a question, such as, what effect do classroom-based mathematics interventions have on mathematics achievement for primary school-aged children? In order to answer this question, researchers need to find any study that has tested the effect of a mathematics intervention in a primary school setting.

Researchers then develop and run a comprehensive search strategy that allow them to find as many relevant studies as possible. The search strategy includes the databases or sources that are searched and the search terms that are used. Searches should also look for relevant unpublished research, such as Master's or PhD theses, which have not been published in academic journals. This helps to avoid publication bias or the tendency of the academic community to publish studies that show positive intervention effects rather than those that show negative or no effects.

Studies identified through the search strategy are then collated together and checked against the review inclusion criteria. Inclusion criteria may relate to the type of intervention, the setting in which the intervention is delivered, the nature of study participants, or the way the research was designed.

Importantly, everything that happens as part of a systematic review, from the search strategy through to the methods of analysis, is decided in advance, and ideally, written up as a review protocol that is published online (for example through (**https://www.crd.york.ac.uk/prospero/**). This makes the systematic review process as transparent as possible thus avoiding bias.

The aims of this review

By systematically reviewing evidence on the effectiveness of a range of mainstream student-directed classroom-based interventions we aimed to allow teachers and policy makers access to reliable and validated research to make informed choices about appropriate teaching practices and interventions. The included trials were classroom-based mathematics interventions directed at primary school-aged pupils, such as those applied at the whole-class level and one-to-one support, delivered by trained teaching professionals as well as through other mediums such as peer tutoring.

This wide variety of potential interventions to improve mathematical achievement may lead to disparate messages being communicated to educators about effective strategies to improve learning outcomes. Given that there is little evidence of transfer of training from domain-general interventions, such as working memory training, to mathematical achievement, or in fact some reports of negative impact (e.g. Roberts et al., 2016), interventions with mathematical content are regarded as a more probable candidate for successful outcomes.

Therefore, this broad review of interventions that aim to improve mathematical learning for children is timely due to the rapid development of new interventions. An evidence base that includes a broad, rather than narrow, review of intervention literature may benefit all children in the classroom.

This systematic review aimed to answer the following questions:

- What types of classroom-based interventions that have been assessed through studies are used with primary school-aged children who do not meet the criteria for mathematics disability?
- What are the most effective classroom based interventions for improving mathematical learning in primary school-aged children who do not meet the criteria for mathematics disability?
- What are the key characteristics of the most effective interventions for improving mathematical learning in primary-school aged children who do not meet the criteria for mathematics disability (e.g. equipment involved, any associated costs)?

What do we mean by intervention?

In this review the term 'intervention' is defined as a deviation from existing teaching practice. Interventions aimed at parents or requiring parental involvement, or those implemented in afterschool or home settings, are beyond the scope of this review and are not included. In this review we do not define interventions solely as activities that are tailored and used with children who are identified as struggling to learn mathematics, rather an intervention is any change in pedagogic approach from existing practice. Unlike previous reviews, we focus on interventions that did not screen participants into the study based on suspected or diagnosed mathematical difficulties, i.e. mathematics achievement below the 25th percentile on standardised mathematical tests. The studies reviewed here investigated changes in pedagogic practice that were used with children of all achievement levels.

This review therefore addresses the need to identify practices that may have the most benefit for large groups of children in the classroom across the achievement spectrum (e.g. those with lower or average mathematical achievement). Eligible study designs are randomised controlled trials (RCT), due to their recognised capability of providing strong evidence of effectiveness (Akobeng, 2005). Quasi-experimental trials, or studies in which there was an experimental and control group, but participants were not randomly assigned to groups, are also included because they offer indications of promising interventions that can be trialled more rigorously in the future.

Methods

The following method for the review was uploaded on the Figshare website (https://bit.ly/2PMgWMt) prior to the review being carried out. A protocol ensures that researchers stringently follow an outlined plan in a systematic review. There were no protocol breaches in this systematic review; the methods were followed as published in the protocol.

To identify all eligible studies, an extensive search was conducted using the following databases:

- British Educational Index
- Education Abstracts
- Academic Search Complete
- MEDLINE
- ERIC
- PsychINFO
- British Educational Index
- Social Sciences Citation Index
- International Bibliography of the Social Sciences
- Applied Social Sciences Index and Abstracts
- Cochrane Central Register of Controlled Trials
- Evidence for Policy Practice Information and Coordinating Centre (EPPI-Centre)

Searching other resources

The following additional resources were also searched:

- ProQuest Dissertations and Thesis
- Dissertation Abstracts International
- Conference Proceedings Citation Index
- Websites of charitable and funding organisations (Nuffield Foundation, National Numeracy Trust, the Education Endowment Foundation and National Science Foundation (USA))
- Government departments (e.g. Department of Education)
- World Health Organisation trials website and clinicaltrials.gov
- Google and Google Scholar (e.g. first 150 hits recorded, exact URL and date of access recorded)
- OpenGrey

The search terms used in the electronic databases are listed in Table 1. These terms are entered into each of the databases above to identify materials that contain specific terms.

In addition, following best practice (Hammerstrøm, Wade, & Jørgensen, 2010), current issues of journals in which included articles had been found were hand searched. Reference lists of all included articles or review articles were checked. Prominent authors in the field were also contacted for details of any recently completed or published studies that may have met the inclusion criteria. No additional studies were identified through this process.

Population	((Primary OR Elementary OR Kindergarten* OR "Grade 1" OR "Grade 2", "Grade 3", "Grade 4", "Grade 5") AND (school* OR educat* OR class* OR teach* OR learn* OR instruct* OR train* OR program*))
Intervention	(Math* OR "Number Sense" OR Numer* OR Arithmetic* OR counting OR addition OR subtraction OR multiplication OR division OR Adding OR Geometry OR fractions OR algebra OR "place value")
Outcomes	(Achieve* OR "Standardi* Test" OR Anxiety OR Attitud* OR "Self-Efficacy" OR Confidence OR Enjoyment)
Methods	(Trial OR RCT OR Quasi OR Random* OR "Control Group" OR "Post Test" OR experimental)

Table 1: Search terms

Criteria for inclusion and exclusion from the search

The following criteria were used to assess if a study would be included in the final systematic review.

Types of studies

Included:

- Randomised controlled trials (including cluster randomised controlled trials)
- Quasi-experimental designs
- Must include comparison control condition (e.g. no intervention, practice-as-usual, waiting list, or active control group)

Excluded:

- Matched subject or group designs
- Cross-over designs
- Single-subject designs
- Correlational designs

Time and language:

- All eligible studies were published from January 2000 to ensure that the materials included were relevant in terms of curriculum context, to the time of literature search conclusion (up to August 2017). Full texts must have been available in English.

Types of participants

Included:

- Participants are children aged 4-11 years' old
- Attending mainstream primary/elementary-level schools

Note: If a study only mentioned Grade level, rather than age, the following criteria were applied: Studies with students between Kindergarten up to Grade 6 were included only if students were situated within a primary/elementary school environment (e.g. middle school Grade 6 were excluded despite including 11-12 year olds).

Excluded:

- Samples selected due to suspected or diagnosed mathematical difficulties (e.g. mathematics achievement below the 25th percentile or less on standardized mathematical tests)
- Samples attending special education establishments
- Samples with other diagnosed learning difficulties or developmental disorders

Types of interventions

Included:

- Interventions (specific deviations from existing practice) aimed at improving mathematics skills (e.g. speeded recall of arithmetic facts, flexible strategy use)
- Interventions carried out in school settings, not laboratory based studies
- Administered individually, within small groups or whole-class based
- Delivery method includes those delivered by teachers, teaching/classroom assistants or other trained professionals
- At least one follow-up at post-test was necessary (within one month)

Excluded:

- Interventions aimed at children screened for or suspected of dyscalculia
- Interventions aimed at parents and requiring parent participation

Types of outcome measures

Primary outcome measures

The primary outcome was mathematics achievement, as measured by:

- Curriculum-based outcome measures (e.g. Key Stage assessments)
- Standardised tests (e.g. Wechsler Individual Achievement Test Numerical Operations or Mathematics Reasoning)
- Cognitive experimental measures of specific mathematics skills (e.g. timed recall of arithmetic facts, flexible strategy use)

Secondary Outcomes

Secondary outcomes included attitudinal or affective measures, e.g. attitude towards mathematics; mathematics anxiety levels; mathematics self-efficacy; confidence in or enjoyment of mathematics. In addition, where available, costs associated with the intervention (e.g. unit costs or costs per student, technology costs) were extracted.

Duration of a follow-up

- Immediate post-test outcomes (up to 30 days post-intervention)
- Longer term follow-up grouped into similar periods (e.g. after 6 months) where available.

Types of settings

Only interventions directed towards students, specifically not curriculum-based changes, in mainstream classrooms were included.

Selection of studies

A research assistant (trained in information retrieval and mentored by Cochrane Ireland and the Campbell Collaboration) conducted the initial search following the pre-registered search strategy (Table 1). Titles and abstracts were checked independently by two of the review team against the inclusion/exclusion criteria. Full texts of potentially eligible studies were located and again screened independently by two of the author team. Any decision disagreements were resolved through discussion and consensus as a team. Reasons for excluding studies were documented.

Overall, the preliminary search identified 10,042 potential papers. Title and abstract screening followed by full text screening applying inclusion and exclusion criteria removed 9,962 studies, leaving a total of 80 included studies.

Findings of the review

RCTs provide more rigorous evidence for the effectiveness of interventions than quasi-experimental trials (i.e. studies where participants were not randomly assigned to groups). Quasi-experimental trials can identify promising interventions that may form the basis for future RCT studies. Although these studies do not have as rigorous methodology the information generated by these studies may assist in informing practice. Therefore, we clearly highlight the evidence from studies using these two types of research methodologies in the following summary sections.

Overview of included studies

The search identified 80 papers that met the inclusion criteria for this systematic review, with 45 RCT and 35 quasi-experimental studies (Figure 2). Overall, these studies included data from 14,198 primary school-aged children (min= 16, max= 1808 per study).

The content of each paper was assessed on its suitability for quantitatively synthesising the outcome data to compare the impact of each intervention (commonly described as meta-analysis). The studies focused on a wide variety of mathematical topics and used a wide range of standardized and experimental outcome measures. In fact, 61 studies used at least one author-generated test as an outcome measure. This makes any comparison of data across studies very difficult. In addition, intervention intensity was diverse across studies ranging from one off sessions to interventions lasting up to two years (Table 3-11). Therefore, we concluded that synthesising the outcome data across studies would not be appropriate. For example, it would not be meaningful to quantitatively compare the effect of a single session of strategy training on algebra problem solving with the effect of several weeks of abacus training on mental arithmetic. Instead, we conducted a narrative synthesis of the included papers.

Studies were classified as belonging to a subtheme; this process was carried out by the lead author and all classifications were discussed with two co-authors to reach agreement. These subthemes were classified as belonging to one of two larger themes that encompassed the studies (Table 2). The first theme focused on different Topic Areas in Mathematics. The second theme focused on differing Methods of Instruction. The subthemes within the Topic Areas in Mathematics included conceptual understanding, magnitudes, basic number skills, practice for fluency and strategy use. The subthemes with the Methods of Instruction theme included the use of manipulatives, providing feedback, technology for engagement and varying delivery contexts.

Figure 2: Summary of included studies

00553-05_UU_A4_Report_v7.indd 12 05/06/2019 15:20

Theme	Subtheme	Studies that...	Number of studies
Topic areas in mathematics	Conceptual understanding	...focus on understanding the principles that underlie a domain, i.e., knowing why something works	14
	Magnitudes	...aim to improve knowledge of how whole or rational numbers fit together in a system	4
	Basic number skills	... target fundamental number skills, such as number sense, quantity comparison or counting skills	7
Methods of instruction	Practice for fluency	...encourage fast access to mathematical facts to aid recall and problem solving	7
	Strategy use	... develop optimal approaches to solve a problem or task	12
	Manipulatives	...use objects as learning aids	5
	Feedback	...change information content or format to improve learning	3
	Technology for engagement	...present information in computer-game format to encourage children's engagement	11
	Delivery contexts	...use of varied approaches to deliver content, such as group work, physical activity or songs	17

Table 2: Summary of identified studies

Conceptual understanding

Fourteen studies focused on improving children's conceptual understanding (Table 3), ten studies were RCTs and four were quasi-experimental trials.

All of these interventions demonstrated a significant increase in performance on outcome measures when compared to a control group. Where clearly specified, age groups included in these studies ranged from 5 to 11 years-old.

Four of the studies used visualization methods (i.e. requiring children to use visual diagrams or schemas) that were either paper or computer-based, to attempt to support children's learning about mathematical concepts, specifically geometry (Al-ebous, 2016), number principles (i.e. primes and composite numbers, Ploger & Hecht, 2009) or mathematical concepts more generally (Rutherford et al., 2014; González-Castro et al., 2014). Both the study by Al-ebous (2016) and Ploger and Hecht (2009) indicated significant improvements in performance in the experimental group when compared to a control group. In contrast, the experimental group in Rutherford et al. (2014) displayed only minimal gains in performance over a one-year period and no gains over a two-year period when compared to a control group. González-Castro et al. (2014) reported that their intervention produced significant improvements in formal and informal skills, but not number facts or calculus, when compared to a control group.

Four studies focused on improving children's understanding of mathematical equivalence by restructuring the format of problem presentation (Hattikudur & Alibali, 2010; McNeil et al., 2012; McNeil, Fyfe & Dunwiddle, 2015; Sophian & Madrid, 2003). These studies reported benefits of teaching conceptual understanding of the equals sign using relational tools, with significant improvement in conceptual understanding and problem solving by experimental groups when compared to controls groups. In addition, Fyfe, DeCaro and Rittle-Johnson (2014) assessed the impact of carrying out some problem solving prior to or following instruction on the relational meaning of the equals sign. This study indicated that children who experienced conceptual training prior to problem solving tasks displayed greater procedural and conceptual knowledge than children who did vice versa.

Four studies assessed whether training conceptual understanding would lead to general improvements in mathematical processing. Paliwal (2013) established that structured training on the complement principle (i.e. that both sides of an equation should balance) led to significant gains in understanding of the principle and its application. Baroody et al. (2016) focused on the impact of teaching the conceptual, rather than simply the procedural, basis for different strategies on arithmetic fluency. This study indicated that training the concept of subtraction by addition significantly improved subtraction fluency and the training of the 'use 10' strategy (i.e. using decomposition with problems bridging 10, such as $9 + 5 = 15 - 1$ as 9 is one less than 10) significantly improved addition fluency. Two studies focused on encouraging children to reflect on basic conceptual understanding (Hung, Sun & Yu, 2015; Kaufman, Delazer, Pohl, Semenza & Dowker, 2005). Hung et al. (2015) manipulated subtraction and addition concept training to focus on simple matching arithmetic problems or more challenging tasks that enabled children to more deeply think about addition and subtraction concepts. The children assigned to the challenging tasks displayed better performance in arithmetic. Kaufman et al. (2005) assigned children to either a control group that was trained in procedural skills or an experimental group that work in small teams to solve conceptual and counting tasks through trial and error. The experimental group displayed significant learning in counting skills and mental calculation in contrast to the control group.

Two studies focused on the conceptual understanding that underpins multiplication. Sophian and Madrid (2003) focused on the concept of many-to-one correspondence by training children to focus on the multiplicative relationship of this concept (e.g. when being asked to place 3 stones in each of 4 buckets). The study aimed to assess if there was transfer to ratio understanding, the study provided limited evidence of any improvements when the experimental group was compared to a control group. Park & Nunes, 2001) assessed the impact of teaching multiplication either by repeated addition or via correspondence (i.e. mapping between a number of sets and the quantities contained in those sets). Results indicated that training that highlighted the concept of correspondence was more effective than repeated addition in improving multiplicative reasoning.

Magnitudes

Four studies targeted magnitude understanding (Table 4), more specifically the understanding of fraction, decimal or whole number magnitudes. All studies indicated positive effects of the interventions.

Both Fazio, Kennedy and Siegler (2016) and Kong (2011) developed computerised interventions to improve fraction understanding. Fazio et al. (2016) used a computerised number line game where children were required to place numbers on a number line, with inbuilt feedback, to improve fraction magnitude knowledge. This led to improvements in fraction knowledge when compared to a control group. Kong (2011) assessed whether the use of an online partitioning tool (i.e. a rectangle that can be divided and regrouped into constituent parts) would aid understanding of fractions. Children in the experimental group displayed better understanding of fraction equivalence and the ability to perform arithmetic with fractions than the control group.

Yoon (2015) compared the impact on mathematical achievement of using a number line versus dynamic hundreds chart for multiplication problem solving. Both intervention groups made significant improvements and there was no difference in performance between intervention groups. Durkin and Rittle-Johnson (2012) assessed magnitude understanding after participants were exposed to correct or incorrect examples on a number line, children exposed to incorrect examples learnt more in terms of procedures and concepts than those exposed to correct examples.

Basic number skills

Seven studies aimed to improve general arithmetic or mathematical achievement by training basic numerical skills (Table 5). Six of the seven studies focused on children in the early years of primary education (Hyde, Khanum & Spelke, 2013; Khanum, Hanif, Spelke, Bertetti & Hyde, 2016; Mascia et al., 2015; Obersteiner, Reiss & Ufer, 2013; Praet & Desoete, 2014). Overall, all studies indicated some positive impact on outcome measures.

Three studies focused on training of specific components of basic numeracy. Sood and Mackey (2015) used a number sense intervention to improve basic numeracy skills, focusing on key concepts such as more/less and part-whole relations. Results were mixed in this study. Praet and Desoete (2014) assessed the impact of counting versus number comparison training, with mixed findings on a number of outcome measures. Khun and Holling (2014) investigated the effect of number sense training when compared to working memory training, in a group of 9 year-olds. This study indicated significant improvements in arithmetic in both groups when compared to controls. However, working memory training also improved mathematical word problem solving whilst number sense training did not.

Three studies specifically investigated the impact of training non-symbolic comparison or addition skills (i.e. the ability to accurately compare or combine two sets of dots) on more general mathematics skills. Hyde et al. (2013) observed that training using non-symbolic quantities, compared to symbolic quantities, led to greater gains in non-symbolic addition skills. This study also reported that children who were trained in non-symbolic comparison showed improvements in symbolic mathematics when compared to a control group. Similarly, Khanum et al. (2016) observed that training in non-symbolic comparison was associated with greater speed in completing symbolic addition tasks and more accurate number estimation skills when compared to a control group. Obersteiner et al. (2013) assessed the impact of symbolic, non-symbolic or combined number comparison skills training. Improvements in the trained skill were observed along with improvements in arithmetic skills for all experimental groups when compared to a control group.

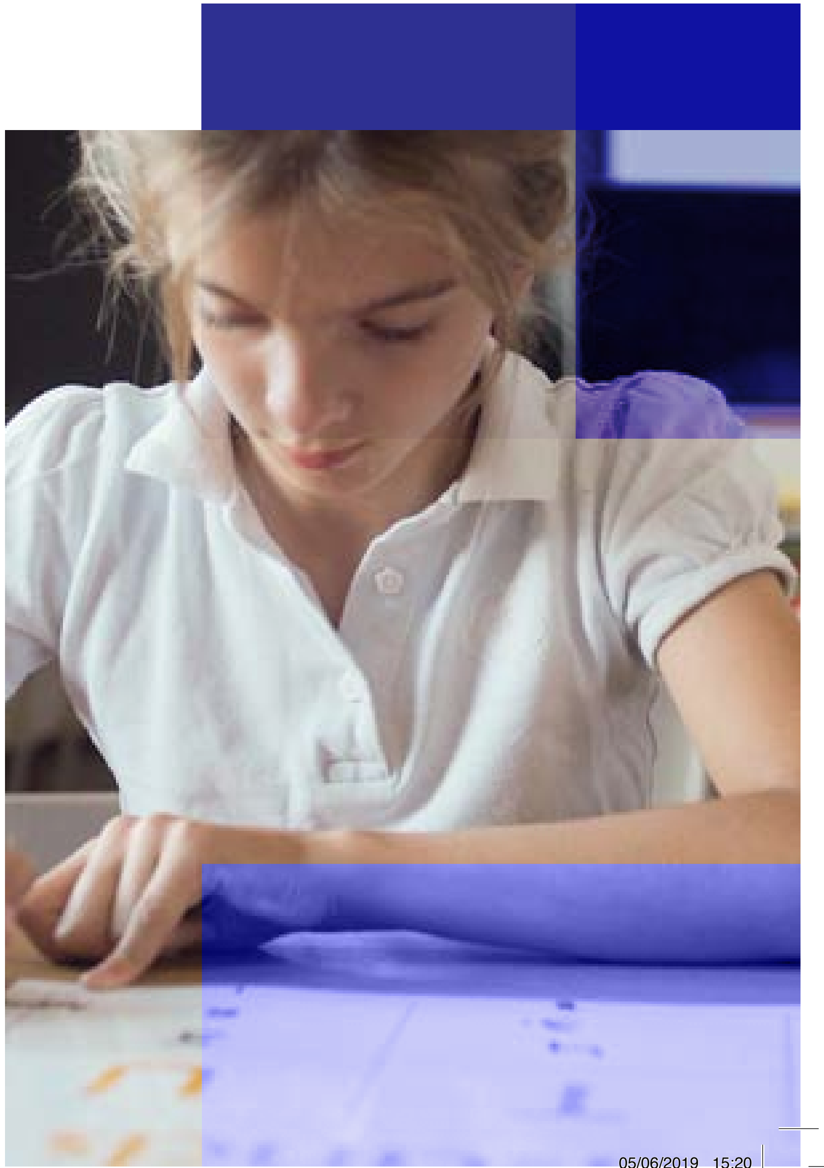
A final study investigated the impact of presentation format (paper and pencil versus computerised tasks) on training basic number skills (Mascia et al., 2015). Improvement in numerical knowledge was observed in both experimental groups when compared to a control group, but there was no difference in performance between experimental groups.

Practice for Fluency

The search identified seven studies that aimed to increase mathematical fluency through a variety of techniques (Table 6). Some mixed results were reported in these studies. Participants in these studies ranged from 5 to 10 years old.

Three studies aimed to increase knowledge of multiplication facts. Burns, Zaslofsky, Maki and Kwong (2015) varied the number of multiplication problems to be learnt using flash cards and observed that tailoring the number of facts to be learnt according to individual children's acquisition rate increased performance. Bakker, van den Heuvel-Panhuize and Robitzsch (2015) observed that computerised mathematics games integrated into mathematics lessons improved children's multiplicative reasoning skills. Zutaut (2002) used mnemonics in order to increase access to multiplication facts. When children using mnemonics were compared to children who were trained through repetition no group differences were observed.

Three studies focused on general arithmetic fluency. Foster, Anthony, Clements and Samara(2016) assessed the impact of a wide variety of games in the Building BlocksTM software that aimed to increase fluency and understanding; generally positive results were observed for the intervention across mathematical problem solving assessments. Cohen (2012) assessed an adaptive computerised program, Mad Dog Math, which trained arithmetic fact retrieval. When compared to a control group a small, but significant improvement was observed in the experimental group in standardised mathematical achievement. Van der Ven, Segers, Takashima and Verhoeven (2017) also used an adaptive computerised program that presented children with numerous symbolic mathematics problems, children were encouraged to focus on speed and accuracy in their responses by being provided with rewards. This study observed that children in the experimental group increased in efficiency in non-symbolic subtraction, but not in symbolic skills, when compared to control children. Caviola, Gerotto and Mammarella (2016) assessed the effect of different practice approaches to mental addition: strategic addition or process based addition (focusing on accuracy and speed). Mixed findings were observed from this, dependent on age. Overall, training improved addition fluency.



Strategy use

Twelve studies intervened on strategy use (Table 7). Prior research has suggested that flexible strategy use and the ability to apply effective and efficient strategies to appropriate problems are related to mathematical achievement (e.g. Geary, Hoard, Byrd-Craven & DeSoto, 2004; Torbeyns, Verschaffel & Ghesquiere, 2004). Eleven of the twelve studies assessed the transfer of strategy training to improvements in mathematical achievement more generally. Overall, eight studies reported positive benefits of the intervention.

Four studies investigated children's problem solving skills and cognitive reflection Gamo, Sander and Richard (2010) assessed the impact of encouraging children to compare similar problems and reflect on the strategies that could be used to solve them. As a follow-up, children were then asked to solve new problems by drawing out their strategy and then reflect on the strategy they used. Children in this experimental group displayed increased strategy efficiency in problem solving compared to a business as usual control group. Mason and Scrivani (2004) focused on improving children's problem solving through use of representations and reflection on strategy use. Both Desoete (2009) and Onu, Eskay, Igbo, Obiyo and Agbo (2012) trained metacognitive, or self-reflection, skills with a specific focus on mathematical content. In addition, Desoete (2009) also encouraged children to reflect on these metacognitive processes through group work. All three studies displayed significant improvements in the experimental condition compared to the control condition.

Three studies focused on problem solving for specific types of mathematical problems. Erkfritz-Gay (2009) assessed the impact of providing and reflecting on visual cues to support strategy use in column addition compared to drill training. No significant group differences in performance were observed. Alibali et al. (2009) assessed the efficacy of training two different algebra problem-solving strategies (equalizing or add/subtract), the equalising strategy led to most substantial improvements in problem solving. Carr, Taasoobshirazi, Stroud and Royer (2011) assessed the efficacy of fluency training versus strategy training for addition and subtraction performance, combined training was also assessed. Combining fluency and strategy training was more effective at increasing mathematical achievement than fluency or strategy training independently. In addition, Jitendra et al. (2007) assessed the impact of training either single or multiple strategy use for mathematical word problem solving. It was observed that single strategy training led to greater gains in mathematical word problem solving and standardised mathematical achievement when compared to multiple strategy training. No differences between groups were observed in computation skills at post-test.

Poland and van Oers (2007) assessed the impact of

00553-05_UU_A4_Report_v7.indd 16 05/06/2019 15:20

schematizing (the use of drawing dynamic diagrams as a strategy for problem solving) on mathematical achievement. Children who experienced schematizing training displayed higher mathematical achievement than a business as usual control group. The impact of drawing schemas (such as diagrams or tables) of mathematical problems was also assessed by Sulak (2010). This study indicated that children who were encouraged to engage in this process developed better problem solving strategies more generally.

Griffin and Jitendra (2008) assessed the impact of schema-based instruction, in which children are encouraged to develop a coherent problem schema to solve a question and then work through this to the solution, on word problem solving. This method was compared to a general strategy instruction group, in which generic step-wise problem solving processes were undertaken. Both groups displayed improvements in problem solving, but improvements were more pronounced in the schema-based instruction group. However, these differences did not persist at an assessment 12 weeks after the intervention. A broad approach was also taken by Hirza, Kusumah and Zulkardi (2014) who compared Realistic Mathematics Instruction (RME) to traditional instruction. RME involves contextualising mathematical learning in real-world scenarios to enable new strategies to be applied to problem solving. Children who experienced RME displayed increased mathematics learning to children taught through traditional methods.

Manipulatives

Five studies focused on the use of manipulatives (physical objects) to improve mathematical performance, all studies reported positive effects (Table 8).

Two studies investigated the efficacy of structured abacus training on more general mathematical skills, such as fluency with mathematics facts and mental arithmetic (Wang et al., 2015; Barner et al., 2016). Both studies began with children using a physical abacus and then transitioned to reliance on mental abacus skills. In an additional study, Gabriel et al. (2012) used physical objects to support teaching fractions; wooden "pie" pieces were used to assist children in estimating fraction addition. The intervention group was reported to have increased performance in fraction knowledge and manipulation.

Two additional studies used virtual manipulatives to improve geometry (Gecu & Staici, 2012) and basic number skills (Pagar, 2013). Gecu and Staici (2012) embedded digital images of real-world scenes into a software tool designed to teach topics such as shape and area. Children were assigned to a group who experienced either a condition in which real-world materials were embedded into the tool or a control condition in which the tool was used by itself. Children who experienced the enhanced training tool displayed significantly higher geometry achievement than the control group. Pagar (2013) assessed the impact of a digital manipulative intervention that either had elements that could be transformed by students or not. Findings were mixed, with the intervention that included the capacity for transformation leading to better counting skills. However, the software, regardless of the ability to transform the materials, led to improved strategy use when compared to an active control group. These two studies indicate the potential importance of manipulation of materials for intervention success.

Feedback

Three studies explored specific types of feedback for learning (Table 9). All feedback was via computerised systems. Two studies used adaptive, personalised feedback that was able to identify specific issues in participants' learning (Chu, Yang, Tseng & Yang, 2014; Faber, Luyten & Visscher, 2017). Both interventions led to positive gains in mathematical achievement when compared to either assessment alone (Chu et al., 2014) or business as usual (Faber et al., 2017). Fyfe and Rittle-Johnson (2016) assessed the impact of computerised feedback either after each task they completed (immediate feedback) or after completing all required tasks (summative feedback) with both conditions displaying significant gains in equation problem solving compared to a control group.

The identified studies indicate that feedback may have differential effects dependent on participant characteristics. Faber et al. (2017) demonstrated that participants with higher performance benefitted more from summative feedback than lower performing participants. In contrast, Fyfe and Rittle-Johnson (2016), noted that only participants with low prior knowledge benefitted from either immediate or summative feedback.

Technology for engagement

Eleven studies used virtual learning environments in order to engage children in mathematical content (Table 10). Children involved in these studies were slightly older than those involved in other themes (8 to 11 years-old). Findings in these studies were mixed.

Five studies assessed the impact of broad computerised training programmes on learning. De Kock and Harkamp (2013) assessed the effectiveness of a web-based problem-solving site that also provided tips on thinking skills. Children who experienced the website significantly outperformed a control group in both problem-solving skills and self-monitoring. Pili and Aksu (2013) assessed the impact of Frisbi Mathematics, computerised mathematics software, on mathematical achievement. Children either experienced the software or were taught traditionally. Children in the experimental group significantly outperformed those in the control group at post-test. Shults (2000) assessed the impact of Math Blaster computerised software. This study found no significant difference in computational skills when children who experienced the software were compared to children who experienced traditional teaching methods. Chang, Sung, Chen and Huang(2008) developed a wide variety of technology-based games focusing on multiplication. It was established that this game only benefitted learning for lower performing children. Rosen and Beck-Hill (2012) assessed the use of computerised one-to-one training to improve mathematics and reading achievement. The computerised software incorporated problem based learning, practice training and game-like tasks. Children who experienced this environment showed significant improvements in achievement compared to a business as usual control group.

Two studies used business environments to engage children in mathematical learning. Barzilai and Blau (2014) established that children needed to study the content that would be practiced in the game first to benefit from this intervention. It is important to note that insufficient information was provided to definitively state that this intervention took place in a classroom environment, but the report stated that the specific web-based game is very popular in schools. Therefore, this study was included. Kim & Ke (2016) used a virtual sandwich bar to teach fraction concepts, significant improvements in mathematical learning were observed for the intervention group compared to the control group.

Two studies assessed whether simply presenting content through digital technology would improve outcomes. Olkun, Smith and Altun (2005) investigated the impact of presenting geometrical puzzle training in a computerised, rather than traditional, format. Children who experienced the computerised training outperformed the children who had experienced the traditional training in geometry understanding. Embedding mathematical instruction in iPad software was found to have no significant benefits for mathematical achievement (as assessed by a standardised measure) in comparison to more traditional methods of content delivery (Singer, 2015).

Huang, Su, Yang and Liou (2016) assessed the impact of the use of a digital pen for teaching mathematics. Children either used the digital pen in small groups, in a lecture theatre set-up or were part of a traditionally taught control group. This study established that children who used the digital pen in groups had significantly better mathematical knowledge than both he children who used the pen in the lecture set-up and the control group. In this study digital technology was not the driver of enhanced learning, rather the group work aspect was the important component.

One study investigated the use of a simple coding programme (in this case Logo) on learning. Yi and EU (2016) studied the effectiveness of teaching knowledge of 2D shapes through Logo. Children were assigned to either a Logo intervention group or a business as usual control group. Children in the experimental group displayed significantly better knowledge of 2D shapes than those in the control group.

Delivery contexts

Seventeen studies focused on the delivery context of mathematical instruction (Table 11). These interventions targeted children throughout the entire primary school age range. Overall, thirteen of the studies produced positive effects.

Four studies focused on the use of group work on learning. One study investigated the influence of group work on children's learning of probability concepts (Gurbuz, Catligou, Birgin & Erdem, 2010). A key component of this intervention was to enable children to question problem-solving strategies, compared to being taught in a didactic manner. Children in the experimental group outperformed children in a control group. Two studies focused on the influence of peer tutoring. Hugger (2012) assessed the impact of peer-assisted learning on mathematical learning and mathematical anxiety. This study also assessed the additive effect of relaxation training on learning. However, at post-test no significant group differences were observed. Tsuei (2012) developed and assessed a peer-tutoring system that enabled children to gain feedback on performance and receive rewards. Children who were assigned to peertutoring, when compared to a control group, displayed significantly higher gains in mathematical achievement and positive self-concept. In addition, the use of multi-touch table tops to teach fractions was compared to the use of computers (Hwang, Shadlev, Tseng & Huang, 2015). Importantly, the table top intervention required children to work in groups. The intervention group showed significant improvements in fraction knowledge when compared to the control group. However, due to the design of the study it is impossible to ascertain if the effect could be attributed to the use of the multi-touch table tops, group-work or the combination of both.

Three studies used movement in their interventions. Mullender-Wijnsma et al. (2015a) used physically active lessons to teach basic arithmetic skills. Movements were related to mathematical content rather than simply using activity to break up sessions. Positive results for the experimental group were observed when compared to a control group. Mullender-Wusma et al. (2015b) embedded physical activity into mathematical lessons. When the experimental group were compared to a traditionally taught control group mixed results were observed. For 6-7 year-olds, the control group significantly outperformed the experimental group in arithmetic skills. The reverse was true for 7-8 year-olds. Ruiter et al. (2015) used number line training either on a ruler on a desk (with or without a mirror to observe their own behaviour) versus a large ruler on the floor. The desk-based tasks required participants to construct double-digit numbers with paper and pencil, the floor-based task required children to indicate the position of numbers by walking to the correct position on the ruler. The full movement condition was associated with the largest improvements in number building.

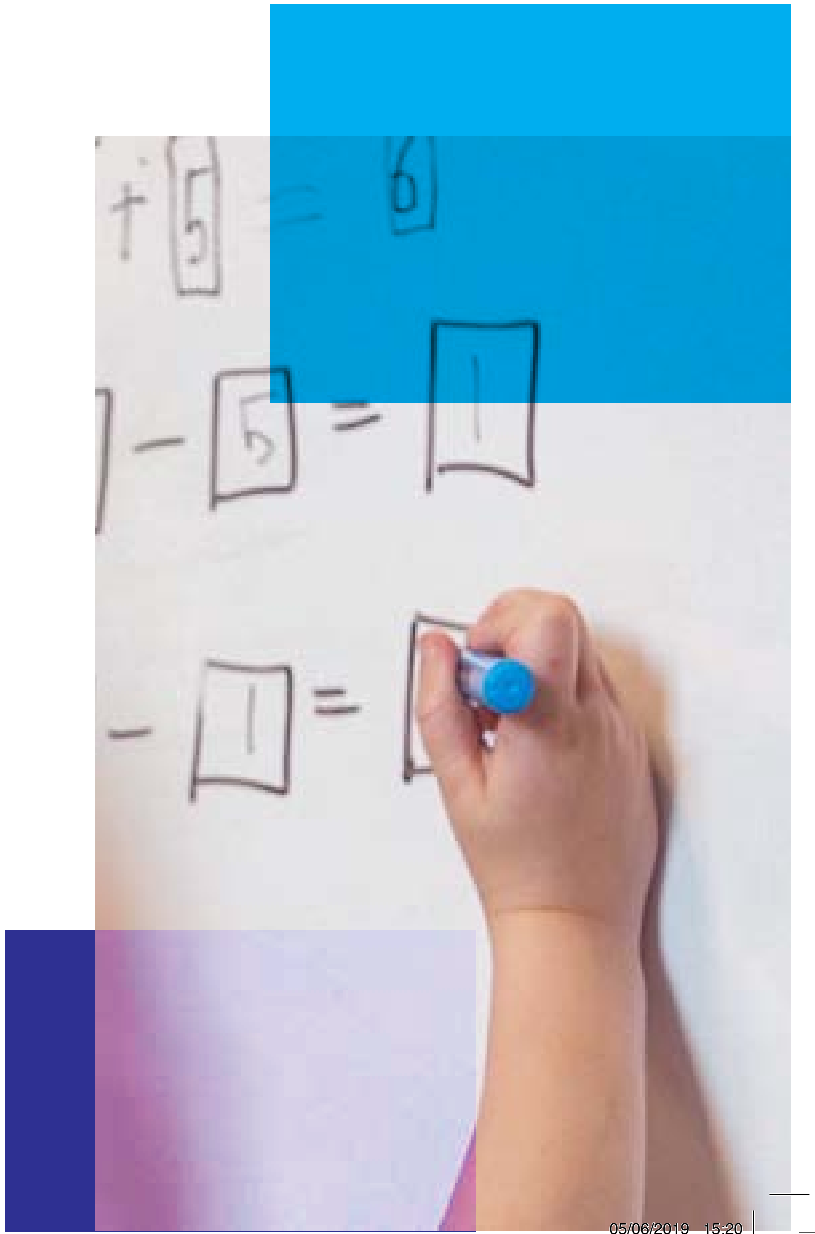
Two studies used music as a delivery method of mathematical information. Courey, Balogh, Siker and Palik (2012) assessed the impact of delivering fraction and concept knowledge through music. This musical intervention focused on the temporal and proportional aspect of music to aid learning. There was no overall effect of the intervention on fraction knowledge when compared to a control group. Kocobas (2009) taught mathematical concepts via mathematical songs. The experimental singing group displayed significantly higher mathematical performance than the control group.

Two studies used story contexts to teach mathematical content to young children. Van den Heuvel-Panhuizen, Elia and Robitzch (2016) assessed the impact of storybook reading on mathematical skills. Casey, Erkut, Ceder and Young (2008) constructed story-contexts to aid teaching geometry concepts. Both studies indicated that story-contexts improved learning when compared to control groups.

Two studies used additional creative approaches to teaching mathematics. Erdogan and Baran (2009) delivered mathematical lessons using techniques from drama, such as role-play. This study indicated that this approach led to significantly higher mathematical achievement than when children were taught via traditional methods, or were taught drama without mathematical content. Csikos and Szitanyi (2012) investigated the effect of children being encouraged to create drawings of mathematical word problems. There was no significant impact of this intervention on arithmetic skills or problem solving when children in the intervention group were compared to the control group.

Three studies undertook different pedagogical approaches to delivering mathematical content. Jeltova, Birney, Fredine, Jarvin, Sternberg and Grigorenko (2011) established that the use of dynamic instruction lead to significantly better geometry and fraction skills when compared to a control group. Dynamic instruction is a pedagogical approach that involves close tracking and responding to children's learning. Lai and Hwang (2016) assessed the effectiveness of flipped classroom learning (i.e. providing children with materials to cover independently before the lesson). This study compared traditional flipped classroom learning to self-regulated flipped classroom learning, in which children set their own learning goals. Children in the self-regulated group displayed significantly higher levels of mathematical learning and better goal setting abilities. Spinner and Fraser (2005) used a Class Banking System (CBS) that focuses on constructivist pedagogical teaching strategies and collaboration between children in classrooms. The application of this approach led to changes in the classroom environment, and higher levels of geometry and algebra knowledge than children taught in traditional classrooms.

An additional study assessed the efficacy of re-studying (re-reading questions with answers provided) versus quizzing (being asked questions and being given feedback) on learning about symmetry (Betsch & Quittenbaum, 2015). Re-studying produced optimal immediate results on performance, in contrast significant positive long-term effects (at 6 weeks follow-up) of quizzing was observed on performance.

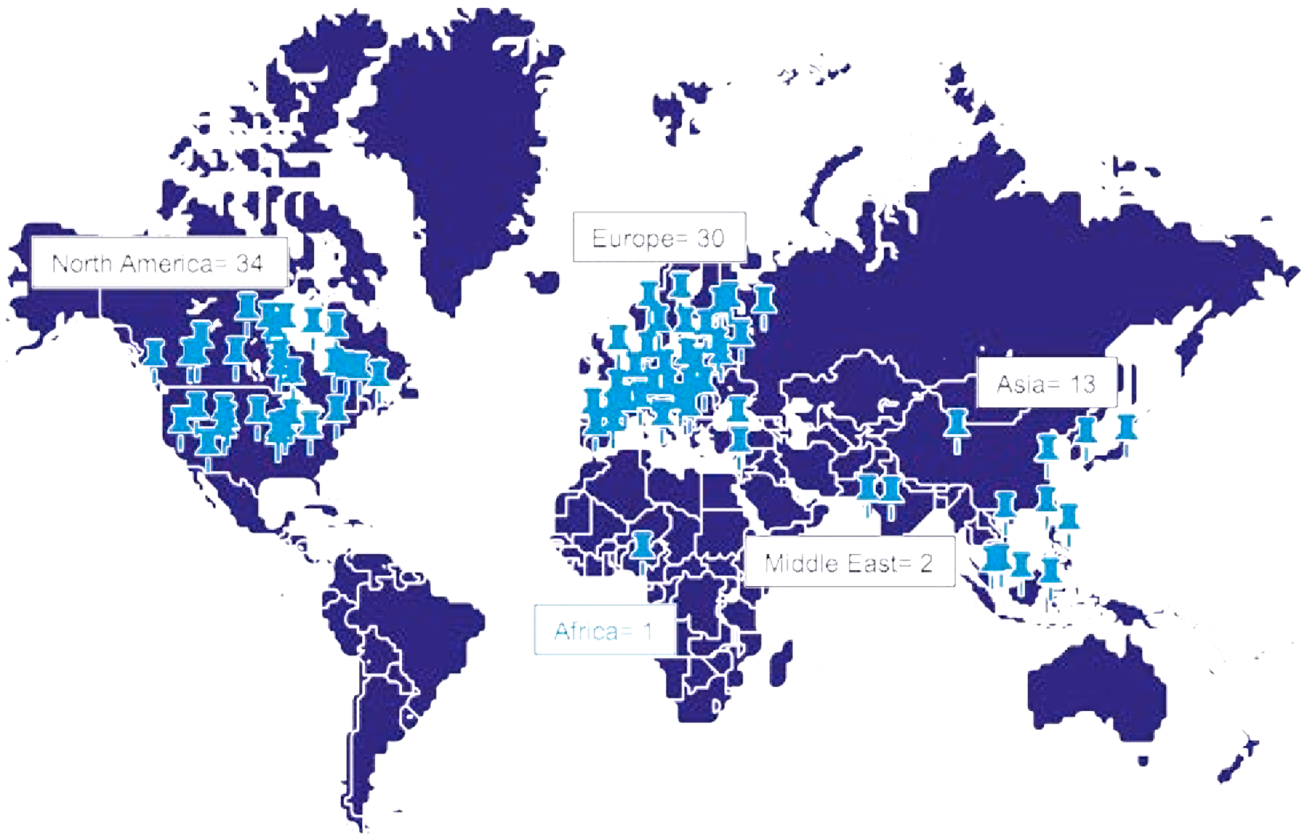


00553-05_UU_A4_Report_v7.indd 19 05/06/2019 15:20

Summary

The narrative summary of the 80 studies clearly displays the wide variety of interventions used to improve mathematical achievement in primary-school-aged children. The majority of interventions reported a positive impact of the assessed intervention on at least one of their outcome measures (72 out of 80 included studies). Interventions were identified from across the globe (e.g. Europe, USA, China and India, Figure 3). The interventions recruited varied sample sizes, ranging from 16 to 1,808 participants. Interventions also varied in intensity from one-off sessions to multiple sessions over a period of two years.

Figure 3: Map depicting geographical spread of included studies



Assessment of rigour and potential for bias in included RCT studies

Even though this report only included studies that used rigorous methods there may still be great variation in the quality of studies that have been identified. In order to assess the rigour and quality of the included RCTs the researchers assessed the potential for bias in each study using **seven recognised criteria** (Cochrane Collaboration's Risk of Bias tool, Higgins & Green, 2011). In order to reduce risk of bias participants should be **(1) randomly assigned to groups using a rigorous and well-documented method**, but **(2) this method should be concealed from participants and researchers whilst data collection is ongoing**. Knowledge of group membership can lead to different expectations for individual's performance, thus biasing trial results. Therefore, risk of bias can also be reduced by ensuring that **(3) participants are not aware of the group that they have been assigned to, (4) individuals delivering the intervention are not aware of other comparison groups in the study** and **(5) researchers assessing outcomes of the study are not aware of group membership of the individuals that they are assessing**. High levels of missing data due to drop-out over the period of the intervention can also lead to biased findings, therefore **(6) clear information should be provided to ascertain that drop out levels are low (i.e. <20%)**. Finally, an additional factor that can influence bias is when authors only report outcomes that have been significantly improved by an intervention. Therefore, studies should be pre-registered. This means that **(7) authors should carefully and publicly outline their research questions and methods prior to data collection**. Then readers can check that all pre-specified measures are reported, thus ensuring that authors have not selectively reported outcome measures.

The overarching issue with all of the included RCT studies was that there was not enough information contained within the papers to make an evidenced-based judgment about the potential for bias across a number of factors. Therefore, it was very difficult to make an assessment of the rigour of the included studies (See Simms, Sloan, McKeaveney & Gilmore, in review, for further details).

Discussion

Summary of results

Through the systematic review process we identified 45 RCT studies and 35 quasi-experimental studies that met inclusion criteria. Given the importance of early mathematical skills for children's current and future life chances (OECD, 2010) it is surprising that so few studies were identified. Practitioners and policy makers therefore have a limited evidence base to inform their practice or policy recommendations.

The wide diversity of the identified studies emphasises a number of issues. First of all, mathematics is a multi-componential subject (Dowker, 2009) and thus the wide variety of target subjects for intervention is not surprising. However, as a consequence of this diversity the literature on mathematical interventions lacks coherence and may not be useful for practitioners or policy makers. Typically, individual studies focus on a very specific outcome and therefore provide little evidence for practitioners seeking to make general decisions about their pedagogical approach in the mathematics classroom. It is also important to note that no replication studies were identified in this review. Replication is essential to ensure that findings are reliable. Therefore even though many studies reported positive findings these have not been replicated in different samples or geographical locations. Thus, the generalizability of these findings must be questioned.

1. What types of classroom-based interventions that have been assessed through studies are used with primary school-aged children who do not meet the criteria for mathematics disability?

The identified studies targeted a number of topics in mathematics, specifically conceptual understanding, magnitudes, basic number skills, mathematics fact fluency and strategy use. Therefore, a diverse battery of outcome measures was used in the included studies. The studies also assessed a variety of different types of delivery methods, specifically the use of physical objects, changes in providing students with feedback, the use of computerised tasks in learning and varying delivery contexts (such as through song or movement). Studies also varied widely in their duration and intensity. Overall, the majority of interventions reported positive results for students learning.

2. What are the most effective classroom-based interventions for improving mathematical learning in primary school-aged children who do not meet the criteria for mathematics disability?

At the protocol registration stage of this systematic review we anticipated that we would be able to use meta-analysis to identify the most effective interventions. However, due to the varied nature and targets of the interventions that the review identified we could not quantitatively compare the studies. For example, it would not be meaningful to try to determine whether strategy training for algebra problems is "more or less effective" than physical objects training for mental arithmetic because they are not attempting to target the same outcome. Instead we identified clear themes that ran across the eighty studies and then qualitatively summarised their approaches and findings. Although quantitative summaries, such as meta-analysis, may appear to simplify the findings across a set of studies with the potential to influence decisions, they make no sense in situations in which interventions and outcome measures vary widely (Eysenck, 1994). Effect sizes cannot be meaningfully compared in this context. This review instead provides a qualitative summary of the included studies that should inform teaching practice.

3. What are the key characteristics of the most effective interventions for improving mathematical learning in primary-school aged children who do not meet the criteria for mathematics disability (e.g. equipment involved, any associated costs)?

Seventy-two out of the eighty included studies (90%) reported a positive impact of the assessed intervention on at least one of their outcome measures. Thus, according to the claims of the study authors the majority of the interventions were effective. However, it is important to note the potential influence of publication bias in this case. It is highly likely that additional trials have investigated interventions that fit our inclusion criteria but may not have been published due to null results. Although the Open Science movement is attempting to rectify the well-known bias in relation to publishing null results, the finding that such a large proportion of included studies reported benefit to learning is striking.

Resource requirements for the studies are summarised in the final column of Tables 3-11 (Figure 6). Only two of the included studies reported the actual cost of the intervention.

Resource requirements for the identified studies

Of the 80 included studies, Fourteen (9%) studies reported that teachers or assistants would require specialist training to deliver the intervention. The majority of interventions did not require expensive equipment for delivery. Three studies required commercially available books to be used as stimuli for learning. Nine studies required basic classroom objects for the intervention (e.g. abacus or playing cards). Twenty-six studies used simple author-generated materials for the intervention (e.g. workbooks and flashcards); teachers could easily produce these materials.

Just over half (N=42) of the studies used computers or tablets as the intervention delivery mechanism. Only two required an interactive whiteboard. Ten studies assessed the efficacy of commercially available software; these products had associated purchasing costs, but the majority were not specified. Twenty-six of the studies assessed the use of author-generated or adapted software.

Some important additional findings

The importance of underpinning theory in developing interventions

When we reviewed each study it was clear that many interventions were based on strong psychological or learning theory. However, there were numerous studies that we identified, in particular summarized within the "Technology for Engagement" subsection (Table 10) that lacked theoretical rationale to explain potential mechanisms for change in outcomes. It is important to recognize that intervention studies can add substantially to pedagogical practice when they test a theory of learning. The results from a study that assesses a theoretically grounded intervention can impact on the development of children's learning more widely. Evidence in support of a theory increases confidence in that theory which may then be generalised beyond the specific context of the study (see for example, Deans for Impact, 2015).

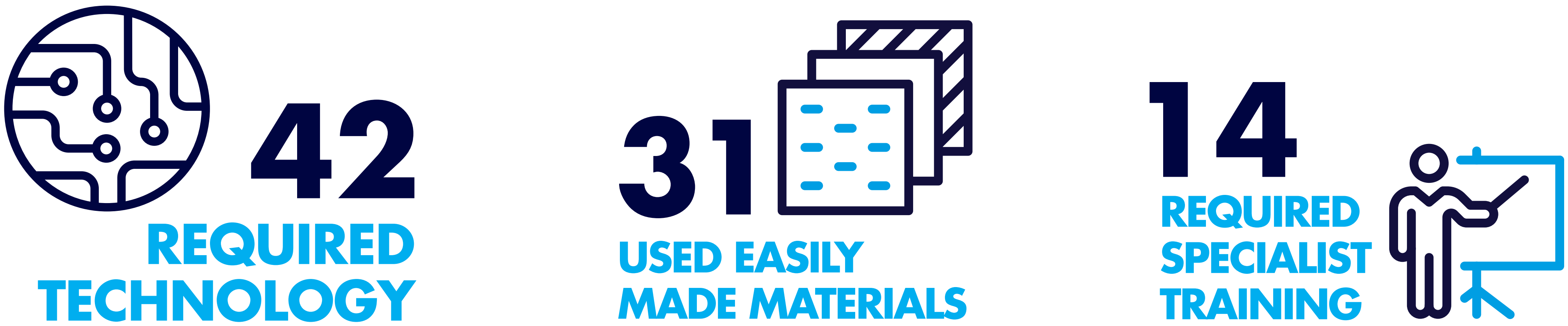
Lack of rigorous studies of commercially available products

The review only identified thirteen studies of commercially available products- these were all either book or software-based interventions. There are numerous commercially available products marketed towards schools, for example those that involve physical objects for teaching mathematical skills. Therefore, the small number of rigorous interventions that were identified in this area was surprising. This review indicates the necessity for rigorous trials of commercially available interventions, such as software and games, to assist teachers and policy makers' decision-making.

Quality of available evidence

An important finding from the review was that many of the studies did not contain sufficient evidence to make an evidence-based judgment about the risk of bias within the study. This is problematic as it is therefore difficult to draw conclusions on the quality of the studies. Thus, for all studies there is the possibility of at least some risk of bias. Many

Figure 6: Summary of resource requirements for included studies



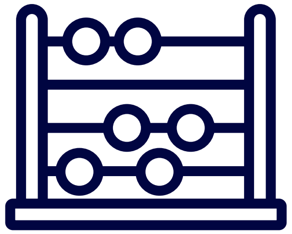
00553-05_UU_A4_Report_v7.indd 22
05/06/2019 15:20

studies included in the review were of low methodological quality when contrasted to the general reporting standards of randomized controlled trials (e.g. http://www.consort-statement.org/). Therefore, we recommend that practitioners and policy makers use the information from this review with caution.

The majority of the studies (N=61) included at least one author-generated outcome measure, which made comparing across studies substantially more difficult. We therefore suggest that researchers, in collaboration with educators, should create a Core Outcome Set that identifies and agrees a standardized set of outcome measurements that can be used across studies, enabling comparison of intervention effects. This process has been commonly used in clinical trials and is supported by programs such as the COMET Initiative (http://www.comet-initiative.org). This does not mean that outcomes are restricted to only those specified in the Core Outcomes Set, but that these outcomes are used alongside pertinent measures for specific studies. Additional measures may include author-generated tasks that, in combination with standardized agreed measures, increase the reach and application of the study findings.

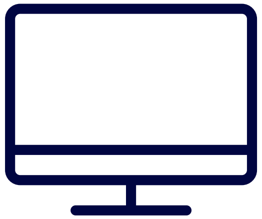
Access to available evidence

All studies were published in academic journals or dissertations and these do not specifically target teaching professionals. The majority (N= 50) were only accessible if the reader paid a fee to the publisher (either independently or through a library subscription). Thus it is clear that many studies may not have impact on teaching practice due to the manner in which they are published. In addition, many studies identified by the review process aimed to test a theoretical framework of learning or mathematical cognition. They were not solely designed to change teaching practice. This of course can be a strength because the evidence these studies produce may have wider implications for theoretical understanding of development and learning. However, the practical applications of these studies may be more difficult to identify.



9

USED EVERYDAY CLASSROOM OBJECTS



10

USED COMMERCIAL SOFTWARE

Limitations of this review

This review focused on interventions that did not target children who were identified as having mathematical learning difficulties. There are numerous existing reviews that focus on interventions to improve mathematical achievement in specific populations of children, for example children with ADHD (DuPaul & Eckert, 1997), emotional and behavioural difficulties (Templeton, Neel & Blood, 2008), in out-of home care (Forsman & Vinnerljung, 2012) and those with mathematical learning difficulties (Dowker, 2009). A consistent finding of these reviews is the efficacy of tutoring for children with specific needs. However, the focus of this review enabled us to identify interventions are appropriate to be used in mainstream classrooms with the objective of improving general mathematical achievement, rather than remediation of mathematical learning difficulties.

This review focused on interventions that are mathematics-specific, i.e. that contain mathematical information in the content. There is a body of evidence supporting a link between mathematical achievement and general thinking skills, such as working memory (Raghubar, Barnes & Hecht, 2010) and visuo-spatial processing (Mix & Cheng, 2012). In addition, behavioural problems and lack of attention has been observed to negatively impact on learning (Barry, Lyman & Klinger, 2002). This review did not include domain-general or general behavioural interventions. Current evidence suggests that interventions that focus on general thinking skills show performance benefits for trained tasks, but low or no transfer to untrained tasks (e.g. Stojanoski, Lyons, Pearce & Owens, 2018). In contrast, some behavioural interventions have indicated positive impact on academic achievement (e.g. Luiselli, Putnam, Handler & Feinberg, 2010). A future systematic review might usefully focus on interventions that primarily target general cognitive skills or behaviour in the classroom and their impact on mathematical achievement.

00553-05_UU_A4_Report_v7.indd 23 05/06/2019 15:20

The current review focused on child-directed interventions rather than interventions that were implemented at institutional or curriculum levels. The authors focused on these child-targeted interventions because, while it is within a teacher's control to change their own practice, this is not necessarily the case at the systemic level. Therefore, this review has practical applications for individual educational practitioners. However, there are three published Campbell reviews that focus on the impact of systemic change in schools on mathematical achievement (along with other measures). Krowka, Hadd and Marx (2017) noted that the introduction of "No Excuses" charter schools were associated with greater gains in mathematical achievement than traditional state funded schools. A small positive impact on mathematics performance was observed through the introduction of school-based-decision making in middle-income countries (Carr-Hill, Rolleston, Schendel, 2016). Changes in teacher recruitment and training through the Teach for America programme had no effect on mathematical achievement (Turner, Ncube, Turner, Boruch & Ibekwe, 2018). Thus, high level changes have mixed impact on achievement.

Existing relevant reviews

In addition to this systematic review there are three existing reviews that teachers and policy makers may find useful. However, findings from these reviews should also be treated with caution due to limitations.

Fischer, Moeller, Cress, Nuerk (2013) conducted a review and synthesized the data (i.e. carried out meta-analysis) from studies to assess the impact of a variety of interventions aimed at improving mathematical achievement for children from pre-school through to secondary school, including those with mathematical learning difficulties. This review was not pre-registered. The authors made the decision to carry out meta-analysis on a wide variety of outcome measures; in contrast to the decision that we made. Fischer et al. (2013) concluded that efficacy of an intervention was related to the type of intervention and the rigour of the trial. The current review observed that nearly all studies reported positive effects, but the majority of studies had issues with potential of bias.

There are also reviews that focus on interventions that target specific components of mathematics. Codding, Hilt-Panahon, Panahon and Benson (2009) reviewed studies that aimed to improve computation and observed that these trials produced moderate to large effect sizes in mathematical performance. Burns, Codding, Boice & Lukito (2010) observed that acquisition and fluency interventions were particularly effective when they were tailored to children's individual skill levels. As mathematics is a multi-componential subject these specific reviews are useful. However, the current review provides a more general summary of interventions across mathematics topics more broadly.

Implications for research

This systematic review has clear implications for future research. Primarily, the issues with difficulty in comparing across studies due to lack of consistent outcome measures should be addressed by researchers through the introduction of a sector-wide Core Outcomes Set, educational research funders could play an important co-ordinating role in this process. This would assist with synthesis of material, increasing the potential for impact on practice and policy. In addition, assessing the rigour of each study was problematic as no study included sufficient information to make this judgment for all seven key indicators. It is essential that sufficient information is reported in published outputs so that readers can ensure that studies have been rigorously carried out and evaluated.

In addition, this review indicates the importance of pre-registration of trials to ensure transparency in the assessment of educational interventions, especially selective outcome reporting. Pre-registration of rigorously designed studies may also increase the probability of null findings being published, thus increasing the evidence base for mathematical interventions.

No replication studies were identified through the systematic review process. With increasing recognition of the importance of replication to generate bodies of evidence it is apparent that this is urgently required in this topic area. This is particularly problematic in the context that nearly all interventions reported positive gains; these should be reassessed to ensure that these results are replicable. In addition, only a small proportion of these identified studies assessed commercially available interventions (N=13). In order to meaningfully inform educational practice rigorous trials of these readily available teaching tools are required.

Conclusions

This review identified a number of studies that assess mathematical interventions. Overall, these studies suggest that changes in practice can improve mathematical achievement. This can be done through a variety of mechanisms, such as changing the method of delivery of content (e.g. computerised tasks, group-work), providing better and more structured feedback or by using additional tools to aid learning (e.g. abacus, rulers, shapes). In addition, focusing on key skills and concepts (e.g. fluency, strategy use, magnitudes) appears to have wider benefit on general mathematical achievement. However, caution is needed because the methods used in many of these studies allow for some bias in interpretation.

As the studies identified in this review were of varying quality and many are not easily accessible to educational practitioners it is very important that a number of messages are conveyed. We need to communicate to teachers what evidence-based education is, and how to evaluate evidence that may influence their practice. This is particularly pertinent in response to commercially available resources. In addition, we need to ensure that studies are made accessible to teachers, either directly to the study publication or through summary documents, such as this report.

Additional resources

The study team has developed a free to access website that summarises the findings of this review. The web-site has a number of functions: 1) Video footage explaining evidence-based education, systematic reviews and RCTs and the findings of this review, 2) a searchable database of included studies and 3) links to resources for teachers.

The web-link is:
www.ulster.ac.uk/mathsinterventions

00553-05_UU_A4_Report_v7.indd 25 05/06/2019 15:20

00553-05_UU_A4_Report_v7.indd 26 05/06/2019 15:20

Summary tables for included studies

00553-05_UU_A4_Report_v7.indd 27 05/06/2019 15:20

Table 3: Summary of conceptual studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Al-ebous (2016)	Effect of the Van Hiele Model in Geometric Concepts Acquisition: The Attitudes towards Geometry and Learning Transfer Effect of the First Three Grades Students in Jordan.	Conceptual understanding	The impact of using the Van Hiele model to inform the teaching of geometry. This model has five phases: visualisation, analysis, abstraction, deduction and rigour. The control group was taught using conventional methods.	6-8 year-olds	RCT	Not specified	Geometric concepts acquisition test (author generated)	The intervention led to significant improvements with the Van Hiele group outperforming the control group who were taught conventionally.	Jordan	Author developed materials
Baroody, Purpura, Eiland, Reid & Paliwal (2016)	Does fostering reasoning strategies for relatively difficult basic combinations promote transfer by K-3 students?	Conceptual understanding	Assess whether computer-based training that focuses on the conceptual bases for two different reasoning strategies, namely subtraction as addition and use-10, improves mental arithmetic skills. A control drill condition was also used.	4-9 year-olds	RCT	7.5 weeks of preparation for all children then 12 weeks of computerised training. Two 30 min sessions per week.	Mental arithmetic test (author generated)	At post-test the subtraction as addition group outperformed the use-10 and drill group on arithmetic subtraction fluency. The use-10 group outperformed the subtraction as addition and control groups on addition problems.	USA	Author generated computerised tasks, computers

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
González-Castro, Cueli, Cabeza, Alvarez-Garcia & Rodriguez (2014)	Improving basic math skills through integrated dynamic representation strategies.	Conceptual understanding	Assessing the impact of the implementation of an adaptive computerised Integrated Dynamic Representation strategy on informal and formal basic mathematical skills. This intervention involves participants identifying key concepts, visualising these concepts and integrating representations. Control participants experienced traditional teaching methods.	6-8 year-olds	RCT	45 50 minute sessions	Mathematical achievement was assessed using the Test of Early Mathematical Abilities (TEMA-3, Ginsburg & Baroody).	The experimental group displayed significant improvement in informal (numbers, comparison, calculation and informal concepts) and formal skills (conventionalisms and formal concepts). There was no significant improvement in number facts or formal calculus when compared to a business as usual control group.	Spain	Specialist training for the teacher delivering the intervention. Tool software and computers

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Hattikudur & Alibali (2010)	Learning about the equal sign: Does comparing with inequality symbols help?	Conceptual understanding	Assess whether instruction that uses comparison of the equals sign to other relational symbols is better at conveying the relational aspect of the equals sign rather than pure instruction about the equals sign alone. Three groups were assessed, a group that was instructed through comparison of the equals sign to other symbols, a group that was directly instructed about the equals and a control group.	8-10 year-olds	RCT	30 mins	Equivalence problems test (author generated)	Children in the comparing condition displayed greater increases in conceptual understanding when compared with the direct instruction group and control group. They also showed better performance in assessments of knowledge of inequality symbols and inequality problem solving in comparison to the other two groups.	USA	Author generated practice sheets

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
McNeil, Chesney, Matthews, Fyfe, Peterson, Dunwiddie & Wheeler (2012)	It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence.	Conceptual understanding	Assess the impact of the structure of arithmetic fact practice on understanding of mathematical equivalence. Problems were either grouped by equivalent sums, grouped by iterative shared addends or no additional practice	7-9 year-olds	RCT	7 weeks (100 min total)	Tests to assess equation solving, equation encoding and defining the equals sign (author generated, used in prior publications)	The group that used facts organised by equivalent values showed better understanding of math equivalence than the other groups.	USA	Author generated workbook

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
McNeil, Fyfe & Dunwiddle (2015)	Arithmetic practice can be modified to promote understanding of mathematical equivalence.	Conceptual understanding	Assess the effectiveness of modifying practice to teach the understanding of equivalence. The control group completed traditional workbooks (i.e. in operations = answer format; 4 + 3 = _). The intervention group completed modified workbooks (i.e. including operations on the right hand side of the equals sign, the replacement of the sign with relational words and blocking of sums with the same answer).	7-8 year-olds	RCT	One off session	Author-generated tests were used to measure equation solving, and equation encoding skills. In addition, children were assessed on their ability to define the equal sign with an author-generated test. Computational fluency was assessed using a timed addition test (Geary, Bow Thams, Lui & Sielger, 1996) and the Math Computation subscale of the Iowa Test of Basic Skills test.	Children in the intervention group displayed better understanding of equivalence than the control group both at post-test and up to 6 months after the intervention. There was no effect of the intervention on computational fluency or performance on the Math Computation test.	USA	Author generated workbook

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Paliwal (2013)	Fostering fluency with basic addition and subtraction facts.	Conceptual understanding	Assess the effectiveness of computerised structured training on the compliment principle compared to unstructured subtraction practice or a control condition that used structured training on a different mathematical topic.	4-9 year-olds	RCT	12 weeks: 30 minute sessions twice per week	Author generated computational shortcut task to gauge understanding of the complement principle. Children must be able to identify the shortcut to complete the problem accurately and quickly	The structured subtraction group outperformed the other groups in knowledge of the complement principle and use of the principle	USA	Author generated computer software, computers.
Park & Nunes (2001)	The development of the concept of multiplication	Conceptual understanding	Assess the impact of teaching multiplication by repeated addition or through correspondence.	5-6 year-olds	RCT	Two school days (part)	Additive and multiplicative reasoning problems (author generated)	Both groups made significant progress in additive and multiplicative reasoning. The group taught by correspondence made more progress in multiplicative reasoning skills than the repeated addition group.	England	Author generated workbook

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Ploger & Hecht (2009)	Enhancing children's conceptual understanding of mathematics through Chartworld software.	Conceptual understanding	Assessment of the efficacy of Chartworld computerised game that teaches arithmetic and number theory. The software allows children to make visual displays of models that can be manipulated. Experiment 1: Focused on the effect of the intervention on conceptual and procedural knowledge of multiplication and division. Experiment 2: Targeted conceptual and procedural knowledge of prime and composite numbers. In both experiments the control group followed a traditional text-book.	8-9 year-olds	RCT	Experiment 1: 8 lessons of 30 mins. Experiment 2: 6 lessons of 45 mins	Structured interview and written test (author generated). The specific details of these assessments were not recorded.	Experiment 1: The chartworld group displayed significant benefits in terms of the skills assessed by the structured interview and the written test. Experiment 2: The Chartworld group displayed greater gains on the semi structured interviews and written test performance when compared to the control group.	USA	Author generated materials in Chartworld software, computers

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Rutherford, Farkas, Duncan, Burchinal, Kibrick, Graham, Richland, Tran, Schneider, Duran & Martinez (2014)	A randomized trial of an elementary school mathematics software intervention: Spatial-temporal math.	Conceptual understanding	This study assessed the efficacy of Spatio-Temporal Math, a computerised teaching programme. This intervention focused on teaching children the links between mathematical concepts through visualisation. A focus on spatial relationships is also included in the intervention. Participants were also recruited into a business as usual control group.	7-11 year-olds	RCT	Two years, with the intervention children receiving 90 minutes additional mathematics instruction per week.	Mathematical achievement was assessed using the California Standard Tests.	The intervention produced a minimal increase in mathematics achievement over a one-year period for all participants when compared to the control group. There were no significant effects over the two year period.	USA	Professional development training for all teachers delivering the intervention. Spatio-temporal Maths software ($35 per student). Computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Hung, Sun & Yu (2015)	The benefits of a challenge: student motivation and flow experience in tablet-PC-game-based learning.	Conceptual understanding	To assess the impact of an IPad application on learning and applying concepts of addition and subtraction. Children were either assigned to an experimental group (who experienced challenging games) or a control group (who experienced matching games).	8-9 year-olds	Quasi-experimental	The intervention consisted of a 40 minute session.	Self-efficacy for science, self-efficacy for technology, flow experiences and feelings about the games were assessed through a questionnaire adapted from previous studies. Mathematics achievement was assessed by a teacher-generated test.	Children assigned to the experimental group displayed better performance than the control group. The experimental group also reported better flow experience and satisfaction.	Taiwan	This intervention required tablet computers and the Motion Math: Hungry Fish application: https://motionmathgames.com/

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Kaufman, Delazer, Pohl, Semenza & Dowker (2005)	Effects of a specific numeracy educational program in kindergarten children: A pilot study.	Conceptual understanding	This study compared the impact of a numeracy-specific intervention that focused on conceptual understanding and counting principles. The intervention was delivered in small groups and children were given opportunities to learn by trial and error. The control group was active and were trained in procedural skills. This training was unstructured.	5-6 year-olds	Quasi-experimental	15 minutes every other day for a full semester	Arithmetic skills were assessed using the Kaufman Assessment Battery for Children-calculation subtest. In addition an author-generated test was used to assess number processing and calculation	Children in the experimental group showed significant learning in counting sequences and mental calculation.	Austria	Author generated materials were used to deliver the instruction. Teachers were provided with monthly supervision by a researcher.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Sophian & Madrid (2003)	Young Children's Reasoning about Many to One Correspondences.	Conceptual understanding	This study investigated the impact of many-to-one mapping training. Experiment 1 focused on the iterative nature of many-to-one correspondence (i.e. the multiplicative relationship when being told to place 3 flowers in each of 4 vases). The experimental group received training that focused on the relationship between large blocks and sets of smaller blocks (i.e. 3 small blocks for each large block, etc), the control group compare heights of stacks of blocks (i.e. to observe when a stack of small blocks reached the same height as two large blocks). Experiment 2 expanded the sample of 7 year-olds and increased the intensity of the intervention.	5-7 year-olds	Quasi-experimental	One off session to deliver a brief intervention (10-15 minutes)	Children's many-to-one correspondence skills were assessed by an author-generated pictorial ratio completion problem test, in which children had to complete some multiplicative relationship to be the same on the left and right side of an equation.	Overall, these experiments provided limited evidence to suggest that the focus of the intervention aided learning, substantial gains were only observed for 7 year-olds- this was driven by large improvements in a small subset of the participants.	USA	Wooden 2D blocks of varying sizes

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Fyfe, DeCaro, & Rittle-Johnson (2014)	An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge.	Conceptual	Children were assigned to either instruct-solve or solve-instruct intervention groups. In the instruction component children were taught the relational meaning of the equals sign through equations. In the problem solving component children completed standard arithmetic problems and math equivalence problems. Children had to explain how they reached their answer. Children received feedback.	7-9 year-olds	Quasi-experimental	One off session lasting 50 minutes (including outcome assessment)	Author generated mathematics equivalence test, adapted from previous publications. This test focused on the meaning of the equals sign and equation structure.	Children who received conceptual training before problem solving led to higher procedural and conceptual knowledge than vice versa. Although, the solve-instruct group did also improve with time.	USA	Author generated problems were generated to enable problem solving.

Table 4: Summary of magnitude studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Durkin & Rittle-Johnson (2012)	The effectiveness of using incorrect examples to support learning about decimal magnitude.	Magnitude	Assessing the efficacy of comparing incorrect and correct examples on learning about decimal magnitude, compared to comparing correct examples.	9-11 year-olds	RCT	25 mins	Author generated conceptual and procedural test for completing decimal number line estimation	The incorrect condition helped students learn correct procedures and key concepts more than the correct condition, including reducing misconceptions.	USA	Resource packet- author generated
Fazio, Kennedy & Siegler (2016)	Improving children's knowledge of fraction magnitudes.	Magnitude	Assessing the impact of a computerised fraction game on children's understanding of fraction magnitudes. This game required children to make number line fraction estimations in a game-based environment; children received feedback on their performance. A control group was also recruited. This group completed the number line estimation game without feedback.	9-11 year-olds	RCT	1 off session (< 15 minutes)	Performance was assessed using an author generated number line estimation test, magnitude comparison test and fraction recall test.	Across two studies, evidence suggested that the intervention increased children's number line estimation, magnitude comparison and recall of fraction information performance. No improvement was observed in the control group's performance.	USA	Catch The Monster computerised game- adapted by authors, computers

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Yoon (2015)	The Effects of Digital Tools on Third Graders' Understanding of Concepts and Development of Skills in Multiplication.	Magnitude	Assessing the impact of digital intervention tools on multiplication skills and number sense. Participants were assigned to either a virtual number line intervention, a dynamic hundreds chart (i.e. a morphable multiplication chart) or a control reading group.	8-9 year-olds	RCT	4 sessions over 4 weeks	Author generated mathematics equivalence test, adapted from previous publications. This test focused on the meaning of the equals sign and equation structure.	Number skills and understanding were assessed using sub-tests from the mClass: Math standardised test. These sub-tests assessed participants' number pattern knowledge and quantity discrimination skills. A timed multiplication test was also administered. This was author generated with specific questions being based on those from an age appropriate text-book. An author generated task also assessed children's understanding of multiplication from different conceptual models.	USA	MathemAnticsTM software developed at Teachers College, Columbia University. Computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Kong (2011)	An evaluation study of the use of a cognitive tool in a one-to-one classroom for promoting classroom-based dialogic interaction.	Magnitudes	This study assessed the effect of a computerised cognitive support tool on achievement, engagement and perceptions of learning fractions. Specifically the Graphical Partitioning Model was used- this is a computerised rectangle that can be partitioned and regrouped in order to scaffold learning. A business as usual control group was also assessed.	9-10 year-olds	Quasi-experimental	11 teaching sessions totalling 455 minutes	Fraction knowledge was assessed using an author generated fraction test- focusing on fraction equivalence, fraction addition and subtraction. An author generated questionnaire was also administered to assess attitudes.	The cognitive tool increased student engagement. Children in the experimental group outperformed the control group in fraction learning. Children reported positive attitudes towards the tool.	Hong Kong	The Graphical Partitioning Model software was author-generated. Children used this scaffold on a tablet computer.

Table 5: Summary of basic number skills studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Khun & Holling (2014)	Number sense or working memory? The effect of two computer-based trainings on mathematical skills in elementary school.	Basic number skills	Comparing the impact of number sense versus working memory training on arithmetic achievement. The number sense training focused on number line estimation and number comparison. The working memory training focused on spatial working memory skills, such as remembering spatial locations. A control group was also recruited.	8-9 year-olds	RCT	20 mins per day for 15 days	Standardised mathematical achievement was assessed using the DEMAT (Krajewski, Liehm & Schneider, 2004)	Both training groups displayed significant improvement in arithmetic achievement. The working memory group showed increased word problem solving over the number sense group.	Germany	Computerised number sense or working memory training game-author-generated. Computers were also required to deliver the intervention.
Mascia, Agus, Fastame, Penna, Sale & Pessa (2015)	The Development and Empowerment of Mathematical Abilities: The Impact of Pencil and Paper and Computerised Interventions for Preschool Children.	Basic number skills	Comparing the impact of paper and pencil versus computerised training. The training was play-based focusing on Arabic number knowledge, such as sequence understanding and estimation. A control group was also assessed.	5 year-olds	RCT	Computer group: 10 x 30 mins; Paper and pencil: 10 x 1 hour	Numerical knowledge was assessed using the BIN (Numerical Intelligence Scale, Molin, Poli & Lucangeli, 2007) that assesses skills such as linking number words to quantities and reciting number word lists.	Both computerised and paper and pencil training groups showed significant improvement in numerical knowledge over time, these were significantly higher than the control group. There was no significant difference in performance between the two intervention groups.	Italy	Training materials from "L'intelligenza numerical" (Lucangeli, Poli, & Molin 2003) and 'Sviluppare l'intelligenza numerical" (Lucangeli, Poli, & Molin 2010) either in paper and pencil format or in computerized format- author generated

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Obersteiner, Reiss & Ufer (2013)	How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills.	Basic number skills	This study aimed to assess the efficacy of exact, approximate or both exact and approximate number processing computer game on arithmetic achievement. A control group was also assessed.	6 year-olds	RCT	10 training sessions of 30 minutes each over a period of 4 weeks	A variety of basic number skills were assessed using author generated tasks, such as: exact and approximate number processing tasks: subitizing, conceptual subitizing, magnitude comparison and approximate calculating. Items from the Hamburger Rechentest were used to assess arithmetic achievement	Improvement was displayed on the trained skill (i.e. exact or approximate skills) but no cross over was observed between intervention groups. Achievement in arithmetic was higher in the experimental groups than the control group at post-test	Germany	Computerised games based on "The Number Race" (Wilson et al., 2006) open-source software: http://www.unicog.org/main/pages.php?page1⁄4NumberRace

Table 5: Summary of basic number skills studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Praet & Desoete (2014)	Enhancing young children's arithmetic skills through non-intensive, computerised kindergarten interventions: A randomised controlled study.	Basic number skills	Assessing the efficacy of computerised interventions on children's early mathematical skills. Children were assigned to either a counting or number comparison intervention. There was also a control group.	5 years-old	RCT	9 x 25 min session	Standardised mathematical achievement was assessed using the TEDI-math (Calculation subtest) and Kortrijk arithmetic test. Number line estimation was also assessed (author generated).	Children in the number comparison and counting group performed better in arithmetic than the control group. Both intervention groups had a better number knowledge compared to the control group. Children in the counting intervention group did better than those in the number comparison group on calculation. There were no group differences in number line estimation performance at follow up.	Belgium	Computer Assisted Intervention serious games-source not specified, trained professionals to deliver the intervention

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Sood & Mackey (2015)	Examining the effects of number sense instruction on mathematics competence of kindergarten students.	Basic number skills	Assessing the efficacy of number sense and general classroom instruction on mathematical skills. The number sense training intervention followed a structured programme focusing on key topics, such as more and less, part-whole relationships, etc. A business as usual control group was also assessed.	Mean age= 5.6 years-old	RCT	All children received 60 minutes of mathematics instruction 4 weeks for 20 mins 5 days a week	Standardised mathematical achievement was assessed using the Stanford Achievement Test (SESAT) subtest of the Stanford Achievement Test (SAT-10). The Early Numeracy-Curriculum Based Measures (EN-CBM) assessed early numeracy skills. Number sense was assessed with an author-generated test.	Both groups made significant gains. For the combined group on the EN-CBM significant improvement for 3 of the 6 sub-tests was reported. There were also significant improvements on all of the 5 subtests on the Number Sense test. The combined intervention group outperformed the control group at delayed post-test on the majority of tests.	USA	Number sense program adapted from existing sources (Bell, Bell, Bretzlauf, Dellard, Hartfeild, Isaacs, et al., 2004; Columba, Kim, & Moe, 2005; Van de Walle, 2007) or author generated

Table 5: Summary of basic number skills studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Hyde, Khanum & Spelke (2013)	Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children.	Basic number skills	In Experiment 1 children experienced either a symbolic addition task, a line-length addition task, a non-symbolic comparison task or a brightness comparison task. Experiment 2 first assigned children to either a non-symbolic addition task or brightness comparison task. Then children were either asked to complete sentence completion tasks or symbolic addition tasks.	6-7 year-olds	Quasi-experimental	The intervention was delivered in a one-off session	In Experiment 1 children's symbolic addition skills were assessed using an author-generated test. Approximate number comparison was assessed using the Panamath computer game. In Experiment 2 participants were asked to complete a sentence completion task or symbolic addition tasks.	In Experiment 1 children who were in either non-symbolic comparison or symbolic addition displayed faster response times for the symbolic addition test. In Experiment 2 children who experienced non-symbolic comparison training were more accurate on the symbolic mathematics problems than children who completed brightness comparison tasks. There was no group difference in sentence completion.	USA	Author-generated computerised tasks and computers.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Khanum, Hanif, Spelke, Bertletti & Hyde (2016)	Effects of non-symbolic approximate number practice on symbolic numerical abilities in Pakistani children.	Basic number skills	Experiment 1 assessed if non-symbolic addition versus line length addition would lead to greater performance in symbolic addition. Experiment 2 assessed if non-symbolic training would transfer to number line estimation performance when compared to control group (line length addition).	6-7 year-olds	Quasi-experimental	One off training session	Experiment 1: Symbolic addition was assessed using an author-generated test. Non-symbolic acuity was assessed using the Panamath task. Experiment 2: Number line estimation was assessed using a paper-based author-generated task.	Experiment 1 indicated that children in the non-symbolic training condition performed symbolic addition faster than those children in the line length condition task. Experiment 2 indicated that number line estimation performance was enhanced in the non-symbolic intervention group compared to a control group.	Pakistan	The intervention was delivered via a computerised task developed by the researchers, computers were therefore required.

Table 6: Summary of practice for fluency studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Bakker, van den Heuvel-Panhuize & Robitzsch (2015)	Effects of playing mathematics computer games on primary school students' multiplicative reasoning ability.	Practice for fluency	Assess the effects of playing mathematics computer games on primary school students' multiplicative reasoning ability. There were four separate groups: mathematics games integrated into lessons, playing games at home with no supervision, playing games at home with feedback and a control group who played other mathematics games on different topics.	6-8 year-olds	RCT	2 years11:29 Children were assessed at the end of both school years	Three components of multiplicative reasoning were assessed: knowledge of number facts, skills in multiplication operations and insight into multiplicative relations. These tasks were author generated.	School based intervention group improved in insight into multiplicative relations only	Netherlands	Computerised game software available for free from: http://www.fisme.science.uu.nl/publicaties/subsets/rekenweb_en/
Burns, Zaslofsky, Maki & Kwong (2015)	Effect of modifying intervention set size with acquisition rate data while practicing single-digit multiplication facts.	Practice for fluency	Assess the impact of varying intervention set size (either 2, 8 or a personalised number of multiplication facts) with acquisition rate (or the amount that a child can learn within a session) data on learning single-digit multiplication facts	8-10 year-olds	RCT	One off session	Multiplication mathematics fact retention was assessed using author generated flash cards. Each fact correctly answered within 2 sec was classified as being recalled.	The tailored acquisition rate group (those who experienced a personalised number of mathematics facts suitable for their acquisition rate) retained significantly more information than comparison groups (who were either taught 2 or 8 mathematics facts).	USA	Flash cards-author generated

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Foster, Anthony, Clements & Samara (2016)	Improving mathematics learning of kindergarten students through computer-assisted instruction.	Practice for fluency	Assess the impact of the Building Blocks Software on mathematical achievement, compared to a computer assisted literacy programme focusing on phonological awareness. The Building Blocks software focuses on developing numeracy and geometry skills, specifically fluency and understanding, through 200 games.	5-7 year-olds	RCT	21 weeks of 90 mins per week	Woodcock Johnson III test of achievement Applied Problems subtest to assess mathematical problem solving and the, Early Maths Assessment (REMA) assessed numeracy skills, with subscales measuring number recognition and subitizing, composition of number, arithmetic and number comparison and sequencing.	The Building Blocks group performed significantly better than the control group on the Applied Problem Solving test and the total REMA test score, the same pattern was observed for all REMA subtests, except for number comparison and sequencing subtest on which the control and experimental group performed similarly.	USA	Building Blocks software. Computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Cohen (2012)	The Effectiveness of Technology Integration in a Metropolitan Elementary Mathematics Program: Mad Dog Math.	Practice for fluency	Assessed the impact of a supplementary program Mad Dog Math intervention that provided additional teaching focusing on retrieval of arithmetic facts. This computerised program is adaptive and provides learners with feedback.	7-9 year-olds	Quasi-experimental	10 weeks	California Standard Test measured mathematical achievement	There was a positive effect of the computerised intervention on standardised mathematics score compared to the control group, although this effect was small.	USA	Mad Dog Math computer https://www.maddogmath.com/ Computers were also required to deliver the intervention.
Van der Ven, Segers, Takashima & Verhoeven (2017)	Effects of a tablet game intervention on simple addition and subtraction fluency in first graders.	Practice for fluency	This study assessed the impact of a tablet game on children's arithmetic fluency. The experimental group played the racing car iPad game, the control group were business as usual. Children were assessed immediately after then intervention and then 13 weeks later.	4-5 year-olds	Quasi-experimental	5 week intervention	Arithmetic skills were measured using an author-generated test, this test was time limited. Tablet game motivation and attitudes towards mathematics were also measured using author-generated questions.	Directly after the intervention the experimental group displayed more efficiency on non-symbolic subtraction than the control group. There was no effect on symbolic efficiency. There was also no significant group effect on motivation	Netherlands	IPads and game software were required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Gamo, Sander & Richard (2010)	Transfer of strategy use by semantic recoding in arithmetic problem solving.	Practice for fluency	The results of two experiments are reported in this paper. In experiment 1 children were assigned to either an experimental or a control group. The control group was business as usual. The experimental group experienced instruction to focus on comparing isomorphic problems and strategies to solve them. On the second day children were presented with new problems and encouraged to draw a representation of the problem, and then compare appropriate strategies. Experiment 2 instructed children to solve the problems using only a single operation, this was not classified as an intervention. No control group was used.	9-11 year-olds	Quasi-experimental	Experiment 1 intervention consisted of two one hour sessions over two days. Experiment 2 did not involve training.	Problem solving skills and strategy use was measured using an author-generated problem solving test.	In Experiment 1, the experimental group displayed increased efficient strategy use compared to the control group. The results of Experiment 2 suggest that the comparison process involved in Experiment 1 may be important for learning.	France	Author generated booklets. Teachers who delivered the intervention received training on problem solving.
Zutaut (2002)	Using Mnemonic Strategies in Fourth Grade Multiplication.	Practice for fluency	This study assessed the effectiveness of training multiplication skills using mnemonics. The experimental group were trained using mnemonics, the control group were trained using repetition.	9-10 year-olds	Quasi-experimental	12 days over a three-week period. Each practice session lasted 10 minutes.	Achievement was measured using an author-generated timed multiplication test.	There were no significant group differences at post-test.	USA	No specific resources were required for this intervention.

Table 7: Summary of strategy use studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Carr, Taasoobshirazi, Stroud & Royer (2011)	Combined fluency and cognitive strategies instruction improves mathematics achievement in early elementary school.	Strategy use	To assess if combined fluency and cognitive strategies instruction improves mathematics achievement in early elementary school. There were four groups in the study: a computerised task targeting increased fluency in addition and subtraction, a program to improve strategy use in addition and subtraction, a combined program and a control group.	7-8 year-olds	RCT	40 half hour sessions over 20 weeks	Cognitive Aptitude Assessment System was used to measure mathematical fluency. An author generated test to assess cognitive strategy use in addition and subtraction problems. Mathematics achievement was assessed using an author-generated test pooling existing available mathematics problems.	There were no significant group differences at post-test.	USA	Computerised training tasks-author-generated. Computers were also required to deliver the intervention.
Erkfritz-Gay (2009)	Differential effects of three computer-assisted instruction programs on the development of math skills among primary grade students	Strategy use	This intervention involved providing examples of the completion of column addition with or without the step-wise strategy visible. There were three conditions: traditional drill and practice, constant time delay or cover copy and compare.	6-7 year-olds	RCT	Daily for three weeks	Author generated mathematics addition computation test to assess fluency and accuracy. Student and teacher acceptability of the intervention was also measured.	No difference in performance between groups at post-test. Participants rated all intervention conditions as acceptable. Teachers rated TDP & CTD better than CCC	USA	Author generated addition problem computerised software. Computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Jitendra, Griffin, Haria, Leh, Adams & Kaduvettoor (2007)	A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving.	Strategy use	Assess the difference in impact of single strategy instruction compared to multiple strategy instruction in improving mathematical problem solving and mathematics achievement.	8-9 year-olds	RCT	25 mins x 5 times a week	Author generated mathematical word problem solving test. The SAT-9 was used to assess mathematics problem solving and procedures. Basic mathematics computation fluency was assessed using a measure by Fuchs, Hamlett & Fuchs (1998). Standardised mathematics achievement was assessed with the Pennsylvania System of School Assessment (PSSA) Mathematics test.	Single strategy instruction led to more favourable gains in mathematical problem solving than multiple strategy instruction, the same pattern was observed for the PSSA standardised test. There were no significant group effects for improvement in computation.	USA	Mathematical word problems-author generated

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Mason & Scrivani (2004)	Enhancing students' mathematical beliefs: An intervention study.	Strategy use	This study aimed to assess the efficacy of cultural change in the classroom towards actively doing mathematics, with mathematical problem solving through representation and staged problem solving. Children were trained to reflect on problem solving strategies as to their appropriateness. A control group was also assessed.	10 year-olds	RCT	12 sessions (1.5 hour per session) over a period of 3 months	An author generated test assessed beliefs about mathematics, mathematics learning and self in relation to mathematics. Mathematical achievement was assessed using an author generated word problem solving. Children reported a self-evaluation of effort. Teachers also provided evaluation of children's mathematics achievement	The experimental group increased advanced beliefs about themselves as mathematics learners, and mathematics and mathematical problem solving when compared to the control group. The experimental group overall outperformed the control group in word problem solving. The control group reported lower perceived effort and understanding than the experimental group.	Italy	Mathematical problem solving tasks- unclear source
Onu, Eskay, Igbo, Obiyo & Agbo (2012)	Effect of Training in Math Metacognitive Strategy on Fractional Achievement of Nigerian Schoolchildren.	Strategy use	To assess the effectiveness of training in math metacognitive strategies following a learning strategy model. The experimental group was compared to a control group.	11 year-olds	RCT	4 weeks of training	Author generated fraction achievement test developed from textbook questions	The experimental group outperformed control group at post-test in achievement in fractional mathematics	Nigeria	Author generated math Metacognitive Training Programme.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Desoete (2009)	Metacognitive prediction and evaluation skills and mathematical learning in third-grade students.	Strategy use	This study assessed the effect of training on metacognition to aid mathematical problem solving. Teachers modelled metacognitive prediction strategies when completing mathematical problems and group work reflected on these processes. A control group was also assessed.	8-10 year-olds	Quasi-experimental	5 times per week for 2 weeks (each session lasted 50 minutes)	Kortrijk Arithmetic Test Revision to measure arithmetic skills, Arithmetic Number Facts Test to measure mathematics fact recall. Cognitive Developmental Arithmetics test assessed arithmetic and metacognition. Metacognitive skills were assessed with Evaluation and Prediction Assessment and a self-report questionnaire.	The experimental group displayed improved metacognitive skills (in term of prediction and evaluation scores). The experimental group also displayed higher arithmetic skills than the control group at post-test and at the beginning of Grade 4.	Belgium	No specific materials were required.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Griffin & Jitendra (2008)	Word problem solving instruction in inclusive third-grade mathematics classrooms.	Strategy use	This study compared performance of problem-solving of children who experienced schema-based instruction (SBI) compared to general strategy instruction (GSI). A key component of the SBI instruction was the development of a problem schema and then working out a problem solution.	8-9 year-olds	Quasi-experimental	The study took place over 20 sessions that lasted 100 minutes each, one day a week. In total the sessions lasted 25 hours over an 18 week period.	Mathematics problem solving and procedures were measured using the Abbreviated Battery of the SAT-9 mathematics test. Mathematical word problem solving was assessed using an author generated word problem solving test. Word problem solving fluency was measured throughout the intervention period using an author generated test. Computation fluency was assessed using a Fluency test (Fuchs & Fuchs, 1998) in which students had to complete as many computations as possible in 3 minutes.	Both groups displayed improved problem solving and computation skills. However, the schema-based instruction group outperformed the general strategy group. The effect of the intervention did not persist to long-term follow-up (12 weeks after the end of the intervention).	USA	Teachers attended two 2 hour long workshops on intervention implementation. Author generated scripts were provided to aid intervention delivery. Author generated one and two-step word problem solving lists. For the SBI condition, teachers were provided with posters displaying schematic diagrams of the three problem types, story and word problem solving check-lists. Students were also provided with workbooks including schematic drawings and word problem solving checklists. For the GSI condition teachers were given posters of word problem steps and students were given manipulatives and workbooks.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Hirza, Kusumah & Zulkardi (2014)	Improving intuition skills with realistic mathematics education.	Strategy use	This study compared Realistic Mathematics Education (RME) to traditional mathematical instruction. RME connects mathematical learning to reality through applied problem solving, applying new strategies. The broad aim was to improve intuition in mathematics problem solving.	11-12 year-olds	Quasi-experimental	The intensity of the intervention is not reported.	Children were assessed on their mathematics intuition ability. The source of this assessment is unclear.	Results indicated that the RME instruction led to greater gains in mathematics learning than traditional instruction.	Indonesia	The required resources for this intervention are not reported.
Sulak (2010)	Effect of problem solving strategies on problem solving achievement in primary school mathematics	Strategy use	This study assessed the impact of problem solving training that focused on strategies compared to a traditional instruction control group.	7-8 year-olds	Quasi-experimental	14 weeks	Children completed two author-generated problem solving tests. These test were administered in the middle and end of the intervention. Qualitative data was also gathered.	Children in the experimental group showed significantly better strategy use than the control group.	Turkey	Resource requirements are unclear.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Poland & van Oers (2007)	Effects of schematising on mathematical development.	Strategy use	This study assessed whether an intervention to encourage schematising (drawing dynamic explanatory diagrams to aid problem solving) would lead to better mathematical achievement 2 years later. Children were assigned to either an experimental schematizing group, who were taught through play-based activities, or a business as usual control group.	5-6 year-olds	Quasi-experimental	Children took part in schematizing activities over the course of a year.	Children's schematizing abilities were assessed using an author-generated test at the end of the intervention period. 8 months after the intervention ended mathematical achievement was assessed using a national standardised test, the CITO.	Children in the schematizing group were observed to have better schematizing skills and higher mathematical achievement.	Netherlands	Teacher-trainer support was provided, the teacher and trainer observed children's activities in order to apply new activities.

Table 8: Summary of manipulatives studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Barner, Alvarez, Sullivan, Brooks, Srinivasan & Frank (2016)	Learning mathematics in a visuospatial format: A randomized, controlled trial of mental abacus instruction.	Manipulatives	Assess whether mental abacus skills can be trained and whether this will influence mathematical abilities. The control group had supplementary text book-based mathematics education.	5-7 year-olds	RCT	3 hours a week over a period of three years	Calculation subtest of the WJ-IIIC, Math Fluency subtest of the WIAT-III and two author generated measures to assess (a) single and multi-digit arithmetic and (b) conceptual understanding of place value	The abacus group showed increased mathematical abilities compared to the control group as measured by the arithmetic and WJ-II Calculation tests, but not on other measures.	India	Abacus, author generated workbooks
Gabriel, Coche, Szucs, Carette, Rey & Content (2012)	Developing children's understanding of fractions: an intervention study.	Manipulatives	Assess the impact of game-like tasks that intended to increase understanding of fractions. The games involved matching and comparing fractions and estimating the outcome of the addition of fractions using objects as support tools. A control group received traditional instruction on fractions.	9-10 year-olds	RCT	Twice a week (30 mins per session) for 10 weeks	Author generated conceptual and procedural test focused on the understanding of fractions	The experimental group scored higher at post-test on conceptual and procedural understanding of fractions when compared to the control group.	Belgium	Playing cards for Memory, War, Old Maid, Treasure Hunt and Blackjack. Wooden disks to represent fractions

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Wang, Geng, Yao, Weng, Hu & Chen (2015)	Abacus training affects math and task switching abilities and modulates their relationships in Chinese children.	Manipulatives	To assess the effectiveness of the use of abacus mental calculation, a traditional method used in Asian countries. First the group learnt with a physical abacus and then move on to using an imagined abacus. Results in the abacus group were compared to a control group.	6 year-olds	RCT	2 hours per week for 26 months (additional assessment taken at 9 months after school entry)	Chinese version of the Heidelberg Rechentest (covers mental addition, subtraction, multiplication, division; number equation completion and number comparison	Abacus training group outperformed the control group on all subscales, significant improvement between grade 2 and 4 for the experimental group not the case for the control group	China	Abacus
Gecu & Staici (2012)	The effects of using digital photographs with Geometer's Sketchpad at 4th grade.	Manipulatives	This study assessed the impact of using digital photographs in combination with a software tool, Geometer's Sketchpad (GSP), to teach geometry (specifically shape, perimeters and areas). These digital images provide real world examples for students to learn with. This software is a virtual manipulatives tool. A control group used the GSP software only.	6 year-olds	RCT	2 hours per week for 26 months (additional assessment taken at 9 months after school entry)	Chinese version of the Heidelberg Rechentest (covers mental addition, subtraction, multiplication, division; number equation completion and number comparison	Abacus training group outperformed the control group on all subscales, significant improvement between grade 2 and 4 for the experimental group not the case for the control group	China	Abacus

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Pagar (2013)	The effects of a grouping by tens manipulative on children's strategy use, base ten understanding and mathematical knowledge.	Manipulatives	This study assessed the effect of a virtual visual manipulative on counting skills, base ten understanding and number sense. Children were either assigned to a Transforming group (children used a computerised manipulative that transformed from a unitized to a continuous presentation), Unitized group (children only viewed computerised unitized models) or a reading software control group.	6-7 year-olds	Quasi-experimental	Children engaged in 5 study sessions, two times per week for 2.5 weeks	Children's ability to identify tens and units was assessed with an author-generated task using concrete manipulatives. Post-test was carried out a week after the end of interventions. Standardised mathematics outcome was measured by mClass: Math Test (Lee et al., 2007). Children also completed a curriculum based measure of basic number skills. Counting on and Base Ten understanding were assessed using measures that were adapted from previous studies.	Children in the Transforming Group displayed significantly better counting on skills at post-test. Both software groups improved in their strategy use and accuracy at post-test. The software interventions improved base ten knowledge for girls only.	USA	Computer software MathemAntics and computers were required to deliver the intervention.

Table 9: Summary of feedback studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Chu, Yang, Tseng & Yang (2014)	Implementation of a Model-Tracing-Based Learning Diagnosis System to Promote Elementary Students' Learning in Mathematics.	Feedback	Assessing the impact of a computerised one-to-one tutoring system. This system was adaptive and identified participants' specific issues with learning mathematics to target feedback. The control group experienced a conventional system that only provided information on accuracy of answers.	11 year-olds	RCT	3 weeks	Author generated paper and pencil fraction knowledge test	The results show that the computerised tutoring system is significantly more helpful to the students in learning mathematics than the conventional web-based test in terms of learning achievements.	Taiwan	Author generated model-tracing Intelligent Tutor (MIT) software- author generated. Computers were also required to deliver the intervention.
Faber, Luyten & Visscher (2017)	The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment	Feedback	This study assessed the impact of an adaptive digital assessment tool (Snappet) that provides formative feedback on mathematical achievement and motivation. A business as usual control group was also assessed.	8-9 year-olds	RCT	5 months	Cito standardised mathematical achievement test (arithmetic, geometry, time and money calculations). The Intrinsic Motivation Inventory was used to measure motivation.	There was a significant and positive effect of the formative assessment tool on mathematical achievement. Higher performing students benefitted more from the formative assessment tool than lower performing students	Netherlands	Snappet formative assessment tool. Tablet computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Fyfe & Rittle-Johnson (2016)	The benefits of computer-generated feedback for mathematics problem solving.	Feedback	An examination of the impact of computerised feedback on mathematical problem solving, assessing the influence of immediate or summative feedback. A control group received no feedback.	7-8 year-olds	RCT	25 mins	Author adapted self-assessment of students' reflection on their performance made them feel. An adapted version of a previously published tool by the authors (Rittle-Johnson et al., 2011) was used to assess equation problem solving.	Immediate feedback had the largest impact on learning. Summative feedback also improved learning compared to no feedback. There was a differential effect of feedback on children with low and high prior knowledge. Children with low prior knowledge benefitted from both immediate and summative feedback, children with high prior knowledge did not benefit from either intervention.	USA	Author generated computerised problem solving task. Computers were also required to deliver the intervention.

Table 10: Summary of technology for engagement studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Barzilai & Blau (2014)	Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences.	Technology for engagement	Test the impact of a business simulation game with additional external conceptual understanding scaffold. There were 3 groups: Study and play (scaffold then game play), Play and study (the reverse) and Play only.	8-11 year-olds	RCT	10 days	Author generated formal problem-solving assessment consisting of mathematical word problems. Perceived learning, enjoyment and flow was also measured.	No significant gains in problem solving were observed. However, the Study and Play group performed better in the post-test problem solving test. The Study and Play condition reduced the participants' perceived learning. The introduction of the scaffold did not impact on enjoyment or flow.	Israel	Games accessed through My Money website: http://money-en.galim.org.il/about.html Computers were also required to deliver the intervention.
Chang, Sung, Chen & Huang (2008)	Learning multiplication through computer-assisted learning activities.	Technology for engagement	Assess the effectiveness of training of concepts, meaning, properties and computation of multiplication in a computerised game environment. Performance in the experimental group was compared to a control group.	8 year-olds	RCT	3 weeks: 120 min teaching sessions per week	Author generated assessment of concepts, meaning, properties and computation of multiplication. Expert feedback to increase validity of the tests.	The intervention improved concepts, meaning, properties and computation of multiplication for lower performing participants only.	Taiwan	Computerised games these were assumed to be author generated. Computers were also required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Kim & Ke (2016)	Effects of game-based learning in an OpenSim-supported virtual environment on mathematical performance.	Technology for engagement	Assess the impact of story game in a virtual reality environment (based on a sandwich bar) on learning of fraction concepts. The control group completed the same tasks on a computer but these were in the form of word problems.	9-10 year-olds	RCT	45 minutes	Author generated fraction test, based on CCSS practice items. Items were verified by teachers for their appropriateness.	Both the experimental and control group improved in their math learning achievement, but the increase was significantly greater for the experimental group.	USA	Author generated virtual Reality environment shop game built in OpenSim software. Computers were also required to deliver the intervention.
Shults (2000)	Teaching First Grade Computation: A Comparison of Traditional Instruction and Computer Enhanced Instruction.	Technology for engagement	This study compared the effect of traditional instruction versus computerised instruction (Math Blaster software) on computational skills.	6-7 year-olds	Quasi-experimental	9 week implementation period. Both the experimental and control group experienced traditional classroom instruction. The experimental group also received an additional 1 hour per week using software, the control group had an additional hour of traditional classroom instruction.	Computation skills were assessed using the computation subtest of the Comprehensive Tests of Basic Skills (4th edition)	No significant group difference in achievement was observed at post-test.	USA	Math Blaster software and computers: http://www.mathblaster.com/

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
de Kock & Harskamp (2013)	Can teachers in primary education implement a metacognitive computer programme for word problem solving in their mathematics classes?	Technology for engagement	To assess the impact of a web-based word problem solving program that included meta-cognitive hints (e.g. to read and analyse, plan, verify). A control group was also assessed.	9-10 year-olds	Quasi-experimental	10 weeks for 20 minutes per week	Log books and questionnaires for teachers to monitor how the intervention was delivered. Log files of each participants performance on the 8 word problem solving tasks. Author-generated analysing word problem test to assess participants' ability to analyse word problems, word problem solving skills test and self-monitoring ability.	The experimental group outperformed the control group on all three outcome measures: analysing word problems, word problem solving and self-monitoring ability.	Netherlands	Teacher workshop to introduce the program and provide examples of how to deliver the material and provide feedback. An author-generated computer program and computers were required to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Huang, Su, Yang & Liou (2016)	A collaborative digital pen learning approach to improving students' learning achievement and motivation in mathematics courses.	Technology for engagement	This study assessed the impact of collaborative learning instruction using a digital pen learning system. Children in the experimental group A were divided into small groups of 4-5 for problem based learning with the digital pen technology. Children in experimental group B used the digital pen in a traditional "lecture" based classroom. A control group was also recruited, this group experienced instruction using traditional pen and paper tasks.	10-11 year-olds	Quasi-experimental	Children experienced the instruction once a week for 4 weeks, each session lasted 120 minutes.	Children's mathematics skills were assessed on mathematics knowledge using a teacher generated test. At pre-test this included 20 multiple choice questions, at post-test an additional 5 open questions were also added to assess problem solving. Children's learning motivation and attitudes to learning were also assessed using previously used questionnaires.	Post-test scores indicated that Experimental Group A had significantly higher post-test scores than Experimental Group B. Experimental Group B also had significantly higher scores than the Control Group. No group differences were observed in learning motivation and attitudes.	Taiwan	Digital pen hardware and software were required. Teacher developed worksheets were also provided.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Olkun, Smith & Altun (2005)	Computers and 2D geometric learning of Turkish fourth and fifth graders.	Technology for engagement	Test the effect of computerised training on geometry learning. Children in the experimental group completed computerised tangram problems. The control group was business as usual.	9-11 year-olds	Quasi-experimental	80-120 minutes per student, dependent on the students need to acclimatise to the software.	Children knowledge of geometry was assessed using an author-generated test.	Children in the experimental group outperformed children in the control group in their geometric test scores.	Turkey	The intervention required computerised 2D tangram problems. Computers were required to run the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Pili & Aksu (2013)	The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus.	Technology for engagement	This study assessed the impact of computerised software on mathematical achievement. A control group was taught using traditional methods. The experimental group was taught using Frizbi Mathematics 4.	9-10 year-olds	Quasi-experimental	Both groups received 4 hours of mathematics instruction following the same mathematics curriculum and books. The experimental group then received an addition 2 hours completed Frizbi Mathematics units. The control group received additional traditional mathematics classroom instruction.	Mathematics achievement was assessed using an author-generated test covering multiplication, division and fractions. Attitudes towards mathematics was assessed using the Mathematics Attitude Scale (MAS, Askar, 1986) and attitudes towards computerised learning was assessed using the Computerised Assisted Learning Attitude Scale (Askar, Yavuz & Koksal, 1991).	Post-test scores were collected at the end of the intervention and 4 months after the intervention. The experimental group outperformed the control group on achievement at post-test and follow-up. This difference was specifically reported for multiplication and division topics, but not fractions. Reported attitudes towards both mathematics and computerised learning in the experimental group were more positive than the control group.	Cyprus	Frizbi Mathematics 4 computerised software was required to deliver the intervention. Computers were also required.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Rosen & Beck-Hill (2012)	Intertwining digital content and a one-to-one laptop environment in teaching and learning: Lessons from the time to know program.	Technology for engagement	This study investigated the impact of computerised one-to-one training on mathematics and reading achievement. The intervention included tailored discovery learning, practice environments and games. Control schools were practice as usual.	9-11 year-olds	Quasi-experimental	90 minutes a day was devoted to both mathematics and English language arts over a school year (6 months)	Achievement was assessed using the Math and Reading Texas Assessment of Knowledge and Skills test. The study also collected data for unexcused absences. Children were asked to complete an author-generated questionnaire on learning motivation and attitudes towards computerised learning.	Significant improvements in reading and mathematics were observed in the experimental group compared to the control group. Improvements in unexcused absences were also reported in the experimental group compared to the control group. Children also reported positive effects of the computerised intervention on motivation to learn and attitudes towards computerised tools for learning.	USA	The use of the Time To Learn Program software and laptops to access the software. All teachers involved in the intervention received pedagogical support.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Singer (2015)	The effects of iPad devices on elementary school students' Mathematics achievement and attitudes.	Technology for engagement	This study assessed the impact of iPad use on children's mathematical achievement and attitudes towards mathematics. An experimental group experienced instruction with imbedded iPads. A control group received traditional instruction. The course content was identical between groups.	8-9 year-olds	Quasi-experimental	One academic year-consisting of 180 days	Children's attitudes towards mathematics was measured with the Attitudes Towards Mathematics Inventory (Tapia & Marsh, 2004). Mathematics achievement was assessed using Pearson SuccessNet benchmarking tool, which is aligned with the school curriculum.	There were no group differences in achievement or attitudes, although teachers qualitatively reported increased engagement, attitudes and productivity.	USA	IPads were required to deliver this intervention.
Yi & Eu (2016)	Effect of using Logo on pupils' learning in two-dimensional shapes.	Technology for engagement	This study assessed the use of the software program Logo on children's understanding of 2D shapes. Children were assigned to the experimental group (taught through Logo) or a business as usual control group.	10-11 year-olds	Quasi-experimental	2 weeks	Achievement in geometry was measured using a non-identified test. Perception towards Logo was measured using a questionnaire.	Performance in the experimental was significantly better than the control group at post-test. Children also reported having positive perceptions of Logo.	Malaysia	Logo software and computers are required to deliver the intervention. A Logo turtle that could be programmed was also required.

Table 11: Summary of delivery context studies

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Betsch & Quittenbaum (2015)	On the Robustness of the Quizzing Effect under Real Teaching Conditions.	Delivery context	This study assessed the impact of scripted direct instruction on children's learning about symmetry. The restudying group read out answers to questions. The quizzing group recalled the answer and were given the correct feedback if incorrect.	9 year-olds	RCT	One 90 minute lesson	Geometry knowledge was assessed using an author generated geometry test, based on questions from textbooks.	The restudying group displayed enhanced performance at immediate follow-up. The quizzing group displayed enhanced learning at 6 weeks follow up.	Germany	Script for teaching symmetry based on the school curriculum, symmetry problems- both author generated.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Casey, Erkut, Ceder & Young (2008)	Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten.	Delivery context	This study aimed to assess the effectiveness of storytelling contexts to improve children's' geometry skills. There were two studies in this paper: (1) Investigated the impact of a story context that focused on part-whole relations in comparison to a control group who received regular mathematics instruction, (2) Compared a story telling context to teach geometry versus traditional instruction on geometry.	5-6 year-olds	RCT	2 days x 20 min session	For Study 1 learning was assessed through tasks that measured both near and far transfer (puzzles using the same pieces as in the intervention and novel puzzle pieces respectively). A modified Kaufman Assessment Battery for Children (K-ABC) Triangles subtest was used to assess near transfer. Far transfer was assessed using a Tangram task (Schiro, 2000).	Study 1 indicated that the intervention only positively affected girls', but not boys', performance on near transfer tasks. Boys' performance increased independent of group membership. There was no effect on far transfer tasks. Study 2 revealed that children in the story telling group improved in both near and far transfer tasks in comparison to the traditional geometry instruction group.	USA	Activity focused "Tan and the Shape Changer" and "Round the Rug Math: Adventures in Problem Solving" books, 2D and 3D shapes.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Gurbuz, Catligou, Birgin & Erdem (2010)	An Investigation of Fifth Grade Students' Conceptual Development of Probability through Activity Based Instruction: A Quasi-Experimental Study.	Delivery context	This study assessed the impact of group discussions on children's learning. Children engaged in group discussions that questioned problem solving processes with the aim to construct knowledge compared to control children who were taught by a teacher in a traditional, didactic manner. A control group was also assessed.	10-11 year-olds	RCT	4 teaching hours	An author generated Conceptual Development Test tested sample space, probability comparisons and probability of an event reasoning skills.	Higher conceptual scores at post-test were observed for all measures for the experimental group than the control group.	Turkey	Probability problems- source unclear.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Hugger (2012)	Evaluating the effects of Peer-Assisted Learning Strategies (PALS) in mathematics plus an anxiety treatment on achievement in third grade students.	Delivery context	This study assessed the impact of mathematics focused peer assisted learning (PAL) on mathematics performance and anxiety. There were four conditions in this study: PAL, PAL and relaxation, relaxation only and a control group.	8-9 year-olds	RCT	The PAL intervention lasted for 12 weeks, children completed two 30 minute sessions per week. The relaxation group completed 5 minutes of relaxation activities combined with a timed 25 minute mathematics intervention per session. The PALS and relaxation session lasted the same amount of time as the other conditions. The control group completed supplementary mathematics activities.	Mathematics computation scores were measured using the AIMSweb Mathematics Computation Curriculum-Based Measure (M-COMP) and anxiety was assessed using the Revised Children's Manifest Anxiety Scale Second Edition (RCMAS-2)	There were no group differences in post-test mathematics computation scores. There was also no significant group difference in anxiety scores.	USA	Teachers received training on the intervention techniques (lasting 1-2 hours). The PAL training group received a manual and worksheets. The relaxation training teachers were provided with a manual and command cards.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Kocabas (2009)	Using songs in mathematics instruction: Results from pilot application.	Delivery context	To assess the impact of delivery of mathematics instruction using group singing in comparison to a control group.	8-9 year-olds	RCT	Academic term	Participants were assessed using two scales (1) Attitude towards Mathematics (Baykul, 1990) and (2) the Scale of Multiple Intelligences (Kocabas, 2007)	The experimental groups' positive attitude towards mathematics increased more substantially than the control groups'. There were no significant improvements in mathematics achievement.	Turkey	Songs with mathematical content recorded to CD- author generated.
Mullender-Wijnsma, Hartman, de Greef, Doolard, Bosker & Visscher (2015a)	Physically Active Math and	Delivery context	Assessing the efficacy of teaching academic lessons through physical activity. Children in the intervention group experiences physically active lessons, children in the control group experienced practice as usual.	7-9 year-olds	RCT	2 years: 22 weeks per year, 3 times a week for 20-30 min sessions. Half of the time in each session was dedicated to mathematics, the other half was dedicated to spelling.	The Speed Test-Arithmetic assessed timed arithmetic performance. Data from the child academic monitoring system was also accessed, this provides a standardised mathematics performance measure.	Results indicated that participants in the intervention group displayed increased performance in the speeded outcome measure compared to the control participants after 2 years. Intervention participants also displayed significantly more improvement in mathematical achievement than the control group after both 1 and 2 years.	Netherlands	Physically Active Lessons manual-author-generated, computer, interactive white-board.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Ruiter, Loyens & Paas (2015)	Watch your step children! Learning two-digit numbers through mirror-based observation of self-initiated body movements.	Delivery context	Assessing the efficacy of teaching double-digit numbers through movement on a number line. Some participants were assigned to a group in which they observed their own behaviour in a mirror. Control participants were asked to construct double-digit numbers on a ruler with pencil and paper. In a second control children were asked to view a ruler across a room, verbally construct the double-digit number and then walk to the point on the ruler.	7-8 year-olds	RCT	One session	Mathematics performance was assessed using an author generated "Number Building Exercises" test that assessed the participants' ability to build two-digit numbers. Participants were also asked their opinion on the intervention and their views on the tasks.	Children in movement conditions were associated with higher mathematics test performance. There was no difference between the mirror and non-mirror movement group in terms of mathematics outcome. There were no group differences in children's opinions of the interventions.	Netherlands	Large floor ruler

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
van den Heuvel-panhuizen, Elia & Robitzch (2016)	Effects of reading picture books on kindergartners' mathematics performance.	Delivery context	This study assessed the impact of delivery of mathematical content through picture story books on mathematical learning. Teachers were provided with scripts to guide their questioning with participants, these activities were completed over and above traditional mathematics instruction. Control children experienced traditional classroom teaching only.	8-9 year-olds	RCT	Academic term	Participants were assessed using two scales (1) Attitude towards Mathematics (Baykul, 1990) and (2) the Scale of Multiple Intelligences (Kocobas, 2007)	The experimental groups' positive attitude towards mathematics increased more substantially than the control groups'. There were no significant improvements in mathematics achievement.	Turkey	Songs with mathematical content recorded to CD- author generated.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Spinner & Fraser (2005)	Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development.	Delivery context	This study assessed the impact of the Class Banking System (CBS) on mathematical learning. This system assists teachers in using constructivist pedagogical approaches through an innovative program, focusing on conceptual development through collaboration with classmates. Children were assigned to either an experimental or a control group.	10-11 year-olds	Quasi-experimental	The intervention lasted one academic year	The classroom environment was measured using the Individualised Classroom Environment Questionnaire and the Constructivist Learning Environment Survey. Attitudes toward mathematics were assessed using the Test of Mathematics Related Attitudes. An author-generated concept map was also administered, this assessed children's knowledge of geometry, algebra, etc. Additional case studies were also completed.	Change in environment scores were significant for both groups on the majority of subscales. The effect size associated with the change was generally larger for the experimental group compared to the control group. Changes in achievement (as measured by the concept mapping task) in the experimental group that were higher than the control were consistently reported.	USA	As this is a change in approach in teaching no additional resources are required.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Tsuei (2012)	Using synchronous peer tutoring system to promote elementary students' learning in mathematics.	Delivery context	This study assessed the effectiveness of a peer tutoring systems on mathematical learning. Children were either assigned to an experimental group who experienced peer tutoring or a control group. The tutoring systems allow children to receive feedback, peer rating and rewards.	10-11 year-olds	Quasi-experimental	Two school semesters with three 40 minute mathematics sessions per week. The first two sessions per week both the control and experimental group received whole class instruction. The third session, students in the control group worked face-to-face in pairs carrying out collaborative tasks. In the third session the experimental group worked online in pairs.	A web-based assessment tool was used to measure children's mathematical achievement. The questions were based on relevant mathematical textbooks. The Self-concept Scale for Children was also administered. And goal motivation was also assessed using the intrinsic goal orientation items from the Motivated strategies for Learning Questionnaire (Pintrich et al., 1993). A mathematical reasoning test (Kramarski & Zoldan, 2008) was used to assess problem solving during the tutoring session. Qualitative data was also collected.	The experimental group significantly greater gains in mathematical achievement compared to the control group. The experimental group also displayed significantly greater increase in self-concept and intrinsic goal orientation when compared to the control group.	Taiwan	G-Math peer tutoring system was used to develop a massive multiplayer online gaming system, computers were used to deliver the intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Mullender-Wusma, Hartmen, de Greef, Bosker, Doolaard & Visscher (2015b)	Improving academic performance of school age children by physical activity in the classroom: 1 year program evaluation.	Delivery context	This study assessed the effectiveness of combining physical activity into learning activities. The experimental group engaged in physically active lessons. The control group experienced traditional lessons.	6-8 year-olds	Quasi-experimental	The intervention was delivered over a period of one year. 63 lessons were developed, each lesson focused on mathematics for 10-15 minutes and then 10-15 minutes on language problems. The intervention consisted of 3 sessions a week for 21 weeks.	Mathematics ability was assessed using the Tempo-Test-Rekenen (Speed Test Arithmetic). Reading ability was assessed using the Een-Minuut-Test (1 Minute test).	For the 7-8 year-olds, children in the experimental group significantly outperformed the control group in reading and mathematics. For the 6-7 year-olds, the experimental group had significantly lower mathematical skills at post-test than the control group. There was no significant group difference for reading skills.	Netherlands	Specialist teachers delivered the intervention. An interactive whiteboard was used in this intervention. Lesson material such as a manual and lesson content.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Erdogan & Baran (2009)	A Study on the Effect of Mathematics Teaching Provided Through Drama on the Mathematics Ability of Six-Year-Old Children.	Delivery context	Children were assigned to an experimental group in which mathematics was taught through drama (such as role play), focusing on skills such as counting, mapping etc. A placebo control group engaged in activities that were not specifically tailored to develop mathematics skills. A control group attended class as usual.	6-7 year-olds	Quasi-experimental	2 days a week for 12 weeks (45-50 minutes sessions per day)	General Information Form and the Test of Early Mathematics Ability-3	Teaching of mathematical skills via drama led to better performance when compared to both the control groups. The placebo control group outperformed the control group.	Turkey	Teaching program "Mathematics Teaching Program Based on the Drama Method" and relevant materials such as cards, boxes, books, cassettes, beads, blocks, costumes.
Courey, Balogh, Siker & Paik (2012)	Academic music: music instruction to engage third-grade students in learning basic fraction concepts.	Delivery context	Assessed the effect of a music intervention on mathematical achievement, specifically on concepts and knowledge of fractions. The intervention focused specifically on temporal and proportional values of musical notes. A control group was also assessed.	8-9 year-olds	Quasi-experimental	6 weeks: Two 45 minutes sessions a week	Author generated music test (focusing on note identification and adding and subtracting notes), fractions concept test and fractions worksheets (focusing on computation)	There was no overall effect of the intervention on the outcome measures. However, further analyses revealed that children who performed poorly on the post-intervention tests and experienced the intervention benefitted in terms of their performance.	USA	Author generated worksheets were used to facilitate learning. Drum sticks and drum pads (mouse pads).

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Csikos & Szitanyi (2012)	The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students.	Delivery context	The impact of visual representations (through drawing) on word problem solving. Participants in the intervention group made their own drawings to aid problem solving and received feedback. A control group was also assessed.	8-9 year-olds	Quasi-experimental	20 lessons (did not exceed 45 minutes per session) over a 5 weeks period	Author generated tests were as follows: 1. Arithmetic skill test that covered curriculum-based topics. 2. A word problem test and 3. A test that assessed beliefs about mathematics.	Although performance on both the arithmetic skill and word problem test improved over the course of the intervention, there was no difference between the intervention and control group in terms of their post-test performance.	Hungary	Teacher training and support throughout the intervention. Author-generated transparencies displaying visual aids, overhead projector.
Hwang, Shadiev, Tseng & Huang (2015)	Exploring Effects of Multi-Touch Tabletop on Collaborative Fraction Learning and the Relationship of Learning Behaviour and Interaction with Learning Achievement.	Delivery context	This study compared the impact of multi-touch table top computers (which involved collaborative learning) compared to traditional computers on fraction learning.	10-11 year-olds	Quasi-experimental	20 minutes class once a week for three weeks	A teacher designed fraction test was used to assess learning. Engagement with the learning materials was also assessed, specifically in terms of time to complete tasks. Children were also asked their opinions of the system and learning motivation via a questionnaire.	Fraction learning was greater in the table top computer group compared to the traditional computer.	Taiwan	Table top computer and tablet computers were required to deliver this intervention.

Author(s) and publication year	Study title	General focus of the study	Intervention summary	Age Group	Research method	Intensity	Outcome measure(s)	Outcome(s)	Country	Intervention resource requirement(s)
Jeltova, Birney, Fredine, Jarvin, Sternberg & Grigorenko (2011)	Making instruction and assessment responsive to diverse students' progress: Group-administered dynamic assessment in teaching mathematics.	Delivery context	This study compared the effects of dynamic instruction compared to triarchic theory or successful intelligence-control instruction. A business as usual control group was also recruited.	9-10 year-olds	Quasi-experimental	There were four units and each unit consisted of 10 lessons taught across a two week period	Outcomes were assessed via either a group administered dynamic post-test or the same test, but dispersed with filler activities. Tests assessed geometry, measurement and equivalent fractions these tests had been used in previous studies by the authors.	The study indicates that dynamic instruction and group based assessment leads to better outcomes.	USA	Teachers received two days' worth of training prior to the intervention. Detailed contents were provided to teachers.
Lai & Hwang (2016)	A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course.	Delivery context	This study assessed the effectiveness of a flipped classroom approach. Children were either assigned to a self-regulated experimental group or a traditional flipped classroom control group. Crucially children in the experimental group were asked to set their own learning goals.	10-11 year-olds	Quasi-experimental	The intervention was applied over three weeks with at least 100 minutes of learning activities in the classroom with additional homework for the flipped component.	Mathematical learning was assessed with a performance test that was developed by three teachers. Self-efficacy and self-regulation were assessed using a questionnaire.	The experimental group displayed higher scores across measures than the control group. Children in the experimental group also displayed an ability to determine goals for their future learning.	Taiwan	Students were provided with e-book and quiz materials through an online system.

References for included studies

Al-ebous, T. (2016). Effect of the Van Hiele Model in Geometric Concepts Acquisition: The Attitudes towards Geometry and Learning Transfer Effect of the First Three Grades Students in Jordan. International Education Studies, 9, 87-98.

Alibali, M. W., Phillips, K. M., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. Cognitive Development, 24, 89-101.

Bakker, M., van den Heuvel-Panhuizen, M., & Robitzsch, A. (2015). Effects of playing mathematics computer games on primary school students' multiplicative reasoning ability. Contemporary Educational Psychology, 40, 55-71.

Barner, D., Alvarez, G., Sullivan, J., Brooks, N., Srinivasan, M., & Frank, M. C. (2016). Learning mathematics in a visuospatial format: A randomized, controlled trial of mental abacus instruction. Child Development, 87(4), 1146-1158.

Baroody, A. J., Purpura, D. J., Eiland, M. D., Reid, E. E., & Paliwal, V. (2016). Does fostering reasoning strategies for relatively difficult basic combinations promote transfer by K-3 students? Journal of Educational Psychology, 108, 576.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. Computers & Education, 70, 65-79.

Betsch, T., Quittenbaum, N., & Lüders, M. (2015). On the Robustness of the Quizzing Effect under Real Teaching Conditions. Zeitschrift für Pädagogische Psychologie, 29, 109-114.

Burns, M. K., Zaslofsky, A. F., Maki, K. E., & Kwong, E. (2016). Effect of modifying intervention set size with acquisition rate data while practicing single-digit multiplication facts. Assessment for Effective Intervention, 41, 131-140.

Carr, M., Taasoobshirazi, G., Stroud, R., & Royer, J. M. (2011). Combined fluency and cognitive strategies instruction improves mathematics achievement in early elementary school. Contemporary Educational Psychology, 36, 323-333.

Casey, B., Erkut, S., Ceder, I., & Young, J. M. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. Journal of Applied Developmental Psychology, 29, 29-48.

Caviola, S., Gerotto, G., & Mammarella, I. C. (2016). Computer-based training for improving mental calculation in third-and fifth-graders. Acta Psychologica, 171, 118-127.

Chang, K. E., Sung, Y. T., Chen, Y. L., & Huang, L. H. (2008). Learning multiplication through computer-assisted learning activities. Computers in Human Behavior, 24, 2904-2916.

Chu, Y., Yang, H. C., Tseng, S. S., & Yang, C. C. (2014). Implementation of a Model-Tracing-Based Learning Diagnosis System to Promote Elementary Students' Learning in Mathematics. Educational Technology & Society, 17, 347-357.

Cohen, M. (2012). The Effectiveness of Technology Integration in a Metropolitan Elementary Mathematics Program: Mad Dog Math. Northcentral University.

Courey, S. J., Balogh, E., Siker, J. R., & Paik, J. (2012). Academic music: music instruction to engage third-grade students in learning basic fraction concepts. Educational Studies in Mathematics, 81, 251-278.

Csíkos, C., Szitányi, J., & Kelemen, R. (2012). The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students. Educational Studies in Mathematics, 81, 47-65.

de Kock, W. D., & Harskamp, E. G. (2014). Can teachers in primary education implement a metacognitive computer programme for word problem solving in their mathematics classes? Educational Research and Evaluation, 20, 231-250.

Desoete, A. (2009). Metacognitive prediction and evaluation skills and mathematical learning in third-grade students. Educational Research and Evaluation, 15, 435-446

Desoete, A., & Praet, M. (2013). Inclusive mathematics education: the value of a computerized look-ahead approach in kindergarten: a randomized controlled study. Transylvanian Journal of Psychology, 14, 103-119.

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. Learning and Instruction, 22, 206-214.

Erdoğan, S., & Baran, G. (2009). A Study on the Effect of Mathematics Teaching Provided Through Drama on the Mathematics Ability of Six-Year-Old Children. Eurasia Journal of Mathematics, Science & Technology Education, 5, 79-85.

Erkfritz-Gay, K. N. (2009). Differential effects of three computer-assisted instruction programs on the development of math skills among primary grade students (pp. 1-190). Illinois State University.

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. Computers & Education, 106, 83-96.

Fazio, L. K., Kennedy, C. A., & Siegler, R. S. (2016). Improving children’s knowledge of fraction magnitudes. PloS one, 11, e0165243.

Foster, M. E., Anthony, J. L., Clements, D. H., Sarama, J., & Williams, J. M. (2016). Improving mathematics learning of kindergarten students through computer-assisted instruction. Journal for Research in Mathematics Education, 47, 206-232.

Fyfe, E. R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. Journal of Experimental Child Psychology, 147, 140-151.

Fyfe, E. R., DeCaro, M. S., & Rittle Johnson, B. (2014). An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge. British Journal of Educational Psychology, 84, 502-519.

Gabriel, F., Coché, F., Szucs, D., Carette, V., Rey, B., & Content, A. (2012). Developing children's understanding of fractions: an intervention study. Mind, Brain, & Education, 6, 137-146.

Gamo, S., Sander, E., & Richard, J. F. (2010). Transfer of strategy use by semantic recoding in arithmetic problem solving. Learning and Instruction, 20, 400-410.

Gecu, Z., & Satici, A. F. (2012). The effects of using digital photographs with Geometer's Sketchpad at 4th grade. Procedia-Social and Behavioral Sciences, 46, 1956-1960.

González-Castro, P., Cueli, M., Cabeza, L., & Rodríguez, C. (2014). Improving basic math skills through integrated dynamic representation strategies. Psicothema, 26, 378-384.

Griffin, C. & Jitendra, A. (2009). Word problem solving instruction in inclusive third-grade mathematics classrooms. Journal of Educational Research, 102, 187-201.

Gurbuz, R., Catlioglu, H., Birgin, O., & Erdem, E. (2010). An Investigation of Fifth Grade Students' Conceptual Development of Probability through Activity Based Instruction: A Quasi-Experimental Study. Educational Sciences: Theory and Practice, 10, 1053-1068.

Hattikudur, S., & Alibali, M. W. (2010). Learning about the equal sign: Does comparing with inequality symbols help? Journal of Experimental Child Psychology, 107, 15-30.

Hirza, Bonita & Kusumah, Yaya & Darhim, Darhim & Zulkardi, Zulkardi. (2014). Improving intuition skills with realistic mathematics education. Journal on Mathematics Education, 5, 27-34.

Huang, C., Su, A. Yang, S. & Liou, H-H.. (2016). A collaborative digital pen learning approach to improving students' learning achievement and motivation in mathematics courses. Computers & Education. 107, 31-44.

Hugger, K.A. (2014). Evaluating the effects of Peer-Assited Learning Strategies (PALS) in mathematics plus an anxiety treatment on achievement in third grade students. University of Southern Maine.

Hung, C. Y., Sun, J. C. Y., & Yu, P. T. (2015). The benefits of a challenge: student motivation and flow experience in tablet-PC-game-based learning. Interactive Learning Environments, 23, 172-190.

Hwang, W. Y., Shadiev, R., Tseng, C. W., & Huang, Y. M. (2015). Exploring Effects of Multi-Touch Tabletop on Collaborative Fraction Learning and the Relationship of Learning Behavior and Interaction with Learning Achievement. Educational Technology & Society, 18, 459-473.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. Cognition, 131, 92-107.

Jeltova, I., Birney, D., Fredine, N., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. (2011). Making instruction and assessment responsive to diverse students' progress: Group-administered dynamic assessment in teaching mathematics. Journal of learning disabilities, 44, 381-395.

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. Journal of Educational Psychology, 99, 115.

Kaufmann, L., Delazer, M., Pohl, R., Semenza, C., & Dowker, A. (2005). Effects of a specific numeracy educational program in kindergarten children: A pilot study. Educational Research and Evaluation, 11, 405-431.

Khanum, S., Hanif, R., Spelke, E. S., Berteletti, I., & Hyde, D. C. (2016). Effects of non-symbolic approximate number practice on symbolic numerical abilities in Pakistani children. PloS one, 11, e0164436.

Kim, H., & Ke, F. (2017). Effects of game-based learning in an OpenSim-supported virtual environment on mathematical performance. Interactive Learning Environments, 25, 543-557.

Kocabaş, A. (2009). Using songs in mathematics instruction: Results from pilot application. Procedia-Social and Behavioral Sciences, 1, 538-543.

Kong, S. C. (2011). An evaluation study of the use of a cognitive tool in a one-to-one classroom for promoting classroom-based dialogic interaction. Computers & Education, 57, 1851-1864.

Kuhn, J. T., & Holling, H. (2014). Number sense or working memory? The effect of two computer-based trainings on mathematical skills in elementary school. Advances in Cognitive Psychology, 10, 59-67.

Lai, C. L., & Hwang, G. J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. Computers & Education, 100, 126-140.

Mascia, M. L., Agus, M., Fastame, M. C., Penna, M. P., Sale, E., & Pessa, E. (2015). The Development and Empowerment of Mathematical Abilities: The Impact of Pencil and Paper and Computerised Interventions for Preschool Children. International Association for Development of the Information Society.

Mason, L., & Scrivani, L. (2004). Enhancing students' mathematical beliefs: An intervention study. Learning and Instruction, 14, 153-176.

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Wheeler, M. C. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. Journal of Educational Psychology, 104, 1109.

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. Journal of Educational Psychology, 107, 423-463.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2015a). Physically active math and language lessons improve academic achievement: a cluster randomized controlled trial. Pediatrics, 137, 2743.

Mullender Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015b). Improving academic performance of school age children by physical activity in the classroom: 1 year program evaluation. Journal of School Health, 85, 365-371.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. Learning and Instruction, 23, 125-135.

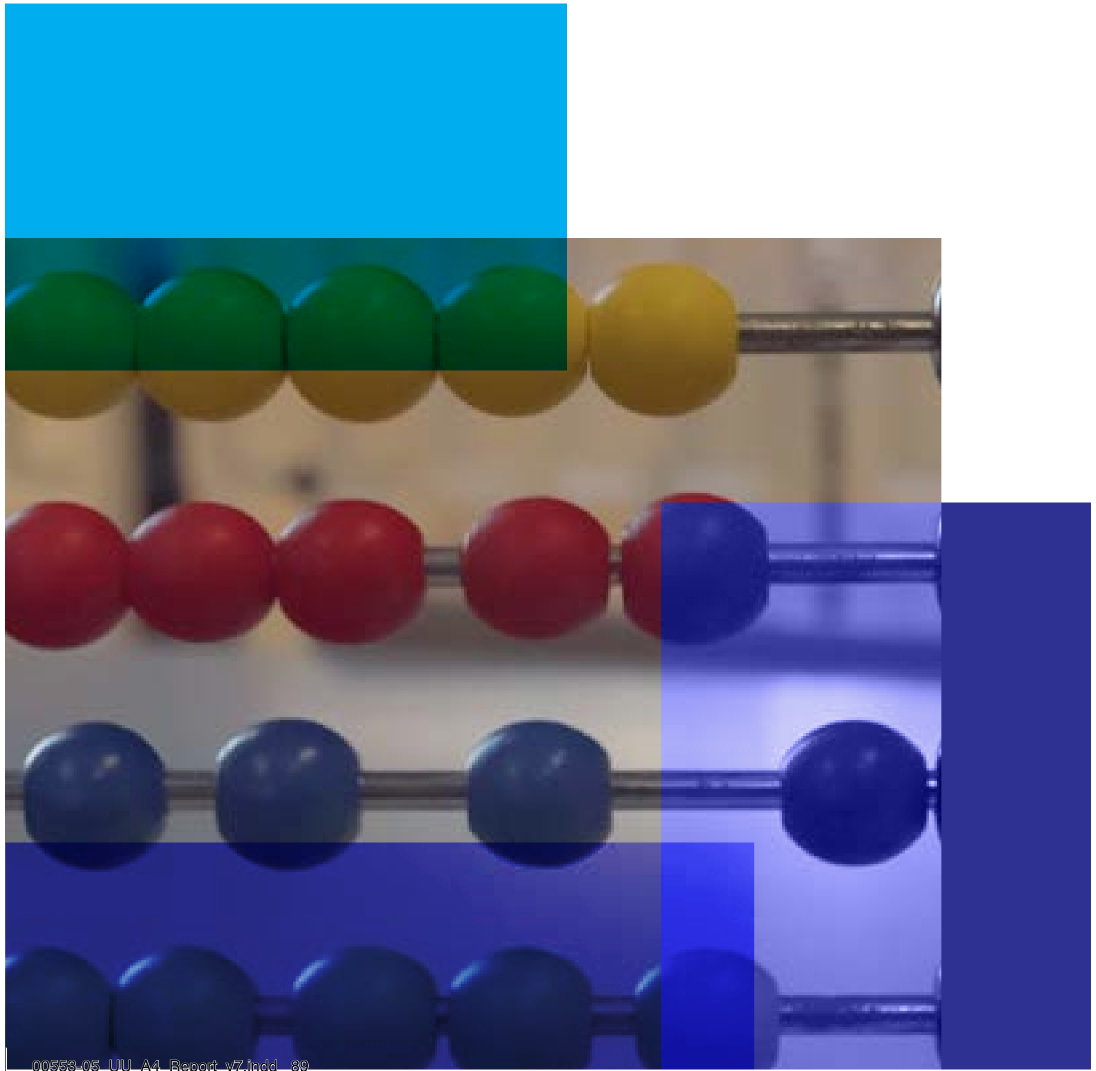
Olkun, S., Altun, A., & Smith, G. (2005). Computers and 2D geometric learning of Turkish fourth and fifth graders. British Journal of Educational Technology, 36, 317-326.

Onu, V. C., Eskay, M., Igbo, J. N., Obiyo, N., & Agbo, O. (2012). Effect of Training in Math Metacognitive Strategy on Fractional Achievement of Nigerian Schoolchildren. US-China Education Review, B3, 316-325.

Pagar, D. (2013). The effects of a grouping by tens manipulative on children's strategy use, base ten understanding and mathematical knowledge. Columbia University.

Paliwal, V. (2013). Fostering fluency with basic addition and subtraction facts. University of Illinois at Urbana-Champaign.

Park, J. H., & Nunes, T. (2001). The development of the concept of multiplication. Cognitive Development, 16, 763-773.



Pilli, O. & Aksu, M. (2013). The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. Computers & Education, 62, 62–71.

Ploger, D., & Hecht, S. (2009). Enhancing children's conceptual understanding of mathematics through Chartworld software. Journal of Research in Childhood Education, 23, 267-277.

Poland, M., & van Oers, B. (2007). Effects of schematising on mathematical development. European Early Childhood Education Research Journal, 15, 269-293.

Praet, M., & Desoete, A. (2014). Enhancing young children's arithmetic skills through non-intensive, computerised kindergarten interventions: A randomised controlled study. Teaching and Teacher Education, 39, 56-65.

Rosen, Y., & Beck-Hill, D. (2012). Intertwining digital content and a one-to-one laptop environment in teaching and learning: Lessons from the time to know program. Journal of Research on Technology in Education, 44, 225-241.

Ruiter, M., Loyens, S., & Paas, F. (2015). Watch your step children! Learning two-digit numbers through mirror-based observation of self-initiated body movements. Educational Psychology Review, 27, 457-474.

Rutherford, T., Farkas, G., Duncan, G., Burchinal, M., Kibrick, M., Graham, J. & Martinez, M. E. (2014). A randomized trial of an elementary school mathematics software intervention: Spatial-temporal math. Journal of Research on Educational Effectiveness, 7, 358-383.

Shults, P. A. (2000). Teaching First Grade Computation: A Comparison of Traditional Instruction and Computer Enhanced Instruction. Johnson Bible College.

Singer, J. (2015). The effects of iPad devices on elementary school students' Mathematics achievement and attitudes. Northeastern University.

Sood, S., & Mackey, M. (2015). Examining the effects of number sense instruction on mathematics competence of kindergarten students. International Journal of Humanities Social Sciences and Education, 2, 14-31.

00553-05_UU_A4_Report_v7.indd 89 05/06/2019 15:20

Sophian, C., & Madrid, S. (2003). Young Children's Reasoning about Many to One Correspondences. Child Development, 74, 1418-1432.

Spinner, H., & Fraser, B. J. (2005). Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development. International Journal of Science and Mathematics Education, 3, 267-293.

Sulak, S. (2010). Effect of problem solving strategies on problem solving achievement in primary school mathematics. Procedia-Social and Behavioral Sciences, 9, 468-472.

Tsuei, M. (2012). Using synchronous peer tutoring system to promote elementary students' learning in mathematics. Computers & Education, 58, 1171-1182.

van den Heuvel-Panhuizen, M., Elia, I., & Robitzsch, A. (2016). Effects of reading picture books on kindergartners' mathematics performance. Educational Psychology, 36, 323-346.

van der Ven, F., Segers, E., Takashima, A., & Verhoeven, L. (2017). Effects of a tablet game intervention on simple addition and subtraction fluency in first graders. Computers in Human Behavior, 72, 200-207.

Wang, C., Geng, F., Yao, Y., Weng, J., Hu, Y., & Chen, F. (2015). Abacus training affects math and task switching abilities and modulates their relationships in Chinese children. PloS one, 10, e0139930.

Yi, B. & Eu, L. (2016). Effect of using Logo on pupuils' learning in two-dimensional shapes. Malaysian Online Journal of Educational Technology, 4, 27-36.

Yoon, E. J. (2015). The Effects of Digital Tools on Third Graders' Understanding of Concepts and Development of Skills in Multiplication. Columbia University Libraries.

Zutaut, A. K. (2002). Using Mnemonic Strategies in Fourth Grade Multiplication. Johnson Bible College.

00553-05_UU_A4_Report_v7.indd 90 05/06/2019 15:20

Additional references

Geary, D., Hoard, M., Byrd-Craven, J. & DeSoto, C. (2004) Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. Journal of Experimental Child Psychology, 88, 121-151.

Torbeyns, J., Verschaffel, L. & Ghesquiere, P. (2004). Strategy development in children with mathematical disabilities. Journal of Learning Disabilities, 37, 119-131.

Akobeng, A. K. (2005). Understanding randomised controlled trials. Archives of Disease in Childhood, 90, 840-844

Barner, D., Alvarez, G., Sullivan, J., Brooks, N., Srinivasan, M. & Frank, M. C. (2016). Learning Mathematics in a Visuospatial Format: A Randomized, Controlled Trial of Mental Abacus Instruction. Child Development, 87, 1146–1158.

Barry, T, Lyman, R. & Klinger, L. (2002). Academic underachievement and attention-deficit/hyperactivity disorder: The negative impact of symptom severity on school performance. Journal of School Psychology, 40, 259-283.

Burns, M., Codding, R., Boice, C. & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. School Psychology Review, 39, 69-83.

Butterworth, B. (2005). The development of arithmetical abilities. Journal of Child Psychology and Psychiatry, 46, 3-18.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. Journal of Educational Psychology, 105, 380-400.

Carr-Hill, R., Rolleston, C. & Schendel, R. (2016). The effects of school-based decision making on educational outcomes in low- and middle-income contexts. Campbell Systematic Reviews, 9. DOI: 10.4073/csr. 2016.9

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

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

COMET Initiative. http://www.comet-initiative.org last accessed on 21st August 2018.

Connolly, P., Keenan, C., & Urbanska, K. (2018). The Trials of Evidence-Based Practice in Education: A Systematic Review of Randomised Controlled Trials in Education Research 1980-2016. Educational Research, 60, 276-291.

Cragg, L. & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. Trends in Neuroscience and Education, 3, 63-68.

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. Cognition, 162, 12-26.

Cowan, R., Donlan, C., Shepherd, D.L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. Journal of Educational Psychology, 103, 786-803.

Davies, P. (2000) The Relevance of Systematic Reviews to Educational Policy and Practice. Oxford Review of Education, 26, 365-378.

Deans for Impact (2015). The Science of Learning. Austin, TX: Deans for Impact.

Department for Education. (2012). Source: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/268031/literacy_and_numeracy_catch_up_strategies_in_secondary_schools.pdf Last Accessed: 31st January 2017.

Department for Education. (2013). Source: https://www.gov.uk/government/news/new-randomised-controlled-trials-will-drive-forward-evidence-based-research Last Accessed: 31st January 2016.

Department for Education. (2015). Source: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/483897/SFR47_2015_text.pdf Last Accessed: 1st January 2016.

Department for Work and Pensions. (2017). Source: https://www.gov.uk/government/publications/employing-disabled-people-and-people-with-health-conditions/employing-disabled-people-and-people-with-health-conditions Last Accessed: 1st February 2017.

Dietrichson, J., Bog, M., Eibeg, M., Filges, T. & Jorgensen, A-M. (2016a). Targeted school-based interventions for improving reading and mathematics for students with, or at risk of, academic difficulties in Grade K to 6: A systematic review. Campbell Systematic Reviews.

Donlan, C., Cowan, R., Newton, E. J., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. Cognition, 103, 23-33

Dowker, A. (2009). What works for children with mathematical difficulties? Nottingham, United Kingdom: Department for Children, Schools and Families. http://www.catchup.org/resources/610/what_works_for_children_with_mathematical_difficulties.pdf

Dowker, A., & Sigley, G. (2010). Targeted interventions for children with arithmetical difficulties. In British Journal of Educational Psychology. Monograph Series II, Number 7-Understanding number development and difficulties (Vol. 65, No. 81, pp. 65-81). British Psychological Society.

DuPaul, G. J. and Eckert, T. L. 1997. School based interventions for children with attention deficit/hyperactivity disorder: A meta analysis. School Psychology Review, 26, 5–27.

Easton, C., & Gee, G. (2012). Early intervention: informing local practice (LGA Research Report). Slough: National Foundation for Educational Research. https://www.nfer.ac.uk/publications/LGLC02/LGLC02.pdf

Every Child a Chance Trust and KPMG. (2009). The long term cost of numeracy difficulties. United Kingdom: Every Child a Chance Trust. https://www.shinetrust.org.uk/wp-content/uploads/ECC-Long-Term-Costs-Numeracy.pdf

Eysenck, H. (1994). Systematic Reviews: Meta-Analysis and its problems. British Medical Journal, 309, 789.

Forsman, H. & Vinnerljung, B. (2012). Interventions aiming to improve school achievements in out-of-home care: A scoping review. Children and Youth Services Review, 34, 1084-1091.

Furlong M, McLoughlin F, McGilloway S, Geary D. Interventions to improve mathematical performance for children with mathematical learning difficulties (MLD) (Protocol). Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD012130. DOI: 10.1002/14651858.CD012130.

Geary, D., & Brown, S. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. Developmental Psychology, 27, 398-406.

Geary, D. C., Bow Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. Child Development, 67, 2022–2044.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. Developmental psychology, 47, 1539.

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. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/wwc/PracticeGuide/2

Gorard, S., See, B.H. & Siddiqui, N. (2017). The trials of evidence-based education. Abingdon, Oxon: Routledge.

Gross, J. (2007). Supporting children with gaps in their mathematical understanding: the impact of the National Numeracy Strategy (NNS) on children who find mathematics difficult. Educational and Child Psychology, 24, 146-156.

00553-05_UU_A4_Report_v7.indd 92 05/06/2019 15:20

Gross, J., C. Hudson, and D. Price. 2009. The Long Term Costs of Numeracy Difficulties. Every Child a Chance Trust and KPMG

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org. Last accessed: 31st January 2017

Kroesbergen, E., & Van Luit, J. (2003). Mathematics interventions for children with special educational needs. Remedial and Special Education, 24, 97-114.

Kormos, J. (2016). The second language learning processes of students with specific learning difficulties. New York: Taylor & Francis.

Krowka, S., Hadd, A. & Marx, R. (2017). 'No excuses' charter schools for increasing math and literacy achievement in primary and secondary education: A systematic review. Campbell Systematic Reviews. DOI: 10.4073/csr.2017.9

Lewis-Presser, A., Clements, M., Ginsburg, H., & Ertle, B. (2015) Big Math for Little Kids: The effectiveness of a preschool and kindergarten mathematics curriculum. Early Education and Development, 26, 399-426.

Luiselli, J., Putnam, R., Handler, M. & Feinberg, A. (2010). Whole-school positive behavior support: effects on student discipline problems and academic performance. Educational Psychology, 25, 183-198.

Maloney, E. A., Schaeffer, M. W., & Beilock, S. L. (2013). Mathematics anxiety and stereotype threat: shared mechanisms, negative consequences and promising interventions. Research in Mathematics Education, 15, 115–128.

Mazzocco, M., (2015) Source: https://cehdvision2020.umn.edu/blog/math-disability-and-difficulty/ Last Accessed: 6th February 2017

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49, 270-291.

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. Journal of Educational Psychology, 10, 423

Mook, D. G. (1983). In defense of external invalidity. American Psychologist, 38, 379.

Morsanyi, K., van Bers, B. M. C. W., McCormack, T. & McGourty, J. (2018). The prevalence of specific learning disorder in mathematics and comorbidity with other developmental disorders in primary school age children. British Journal of Psychology, 109, 917-940.

Mix, K., & Cheng, Y. (2012). The relation between space and math: Developmental and educational implications. In J. Benson (Ed.), Advances in child development and behavior (pp. 197–243). Burlington, VT: Academic Press.

Mulrow, C. (1994). Systematic reviews: Rationale for systematic reviews. British Medical Journal, 309, 597.

Najarian, M., Snow, K., Lennon, J., & Kinsey, S. (2010). Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), preschool–kindergarten 2007 psychometric report (NCES 2010–009). Washington, DC: National Center for Education Statistics. http://www.tandfonline.com/servlet/link?sIT00d=16&doi=10.1080%2F10409289.2015.994451&key=10.1037%2Fe543722010-001

Nelson, J. & Campbell, C. (2017). Evidence-informed practice in education: meanings and applications. Educational Research, 59, 127-135.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. Learning and Instruction, 23, 125-135.

Organisation of Economic Co-operation and Development (2010). Education at a glance. Source: http://www.oecd.org/education/skills-beyond-school/45926093.pdf Last Accessed: 1st January 2016.

Organisation of Economic Co-operation and Development (2016). Education at a glance. OECD indicators. Source: http://www.oecd.org/skills/piaac/Country%20note%20-%20United%20Kingdom.pdf Last Accessed: 1st January 2016.

Raghubar, K., Barnes, M. & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. Learning and Individual Differences, 20, 110-122.

Seethaler, P. & Fuchs, L. (2005). A drop in the bucket: Randomised control trials testing reading and mathematics interventions. Learning Disabilities: Research and Practice, 20, 98-102.

Shadish, W., Cook, T., Campbell, D. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, New York: Houghton Mifflin Company.

Sharples, J. (2013). Evidence for the frontline. The Alliance for Useful Evidence report.

Stojanoski, B., Lyons, K., Pearce, A. & Owen, A. (2018). Targeted training: Converging evidence against the transferable benefits of online brain training on cognitive function. Neuropsychologia, 117, 541.

Templeton, T., Neel, R. & Blood, E. (2008). Meta-analysis of math interventions for students with emotional and behavioural difficulties. Journal of Emotional and Behavioural Disorders, 16, 226-239.

Thurston, A., Tymms, P., Merrell, C. & Conlin, N. Jul 2014 Improving reading and math achievement across a whole district. Proven Programs in Education: Classroom Management and Assessment. USA: Corwin. Chapter 5, p. 116.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence informed management knowledge by means of systematic review. British Journal of Management, 14, 207-222.

Turner, H., Ncube, M., Turner, A., Boruch, R. & Ibekwe, N. (2018). What are the effects of Teach for America on math, English language, arts and science outcomes of K-12 studnets in the USA? Campbell Systematic Reviews, 7. DOI: https://doi.org/10.4073/csr.2018.7

Raghubar, K.P., Barnes, M.A., & Hecht, S.A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. Learning and Individual Differences, 20, 110-122.

Ramirez, G., & Beilock, S. L. (2011). Writing About Testing Worries Boosts Exam Performance in the Classroom. Science, 331, 211-213.

Sarama, J. & Clements, D. H. (2oo9b). Early childhood mathematics education research: learning. New York: Routledge

Sibbald, B. & Roland, M. (1998). Understanding controlled trials. Why are randomised controlled trials important? British Medical Journal, 316, 201.

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

Skills for life (2012) A 2011 survey of literacy, numeracy and ICT levels in England. Department for Business Innovation and Skills. Source: https://www.nationalnumeracy.org.uk/what-issue/ Last Accessed: 27th February 2017

Sweller, J., Clark, R., & Kirschner, P. (2010). Teaching general problem-solving skills is not a substitute for, or a viable addition to, teaching mathematics. Notices of the American Mathematical Society, 57, 1303-1304.

Sykes, Gary, Barbara Schneider, and David N. Plank, (2012) eds. Handbook of education policy research. New York: Routledge.

Wardrop, M. (2014) A Guide to Success with Math: An interactive Approach to understanding and teaching Orthon Gillingham math. Canada: Book Makers.

White, H. (2013). An introduction to the use of randomised control trials to evaluate development interventions. Journal of Development Effectiveness, 5, 30-49.

Wilson, A., Dehaene,S., Dubois,O., & Fayol, M. (2009). Effects of an Adaptive Game Intervention on Accessing Number Sense in Low Socioeconomic Status Kindergarten Children. Mind, Brain, and Education, 3, 224-234.

Xin, Y. P. & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. The Journal of Special Education, 32, 40-78.

Xing, J. (Ed.). (2013). General education and the development of global citizenship in Hong Kong, Taiwan and Mainland China: not merely icing on the cake. New York: Routledge.

00553-05_UU_A4_Report_v7.indd 95 05/06/2019 15:20

Ulster University

@ulsteruni

/ulsteruni

@ulsteruni

ulster.ac.uk

00553-05_UU_A4_Report_v7.indd 96 05/06/2019 15:20