
This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Evidence for review of mathematics teaching: Improving mathematics in key stages two and three

PLEASE CITE THE PUBLISHED VERSION

https://educationendowmentfoundation.org.uk/public/files/Support/Links/Campaigns/Maths/EEF_Maths_Evidenc

PUBLISHER

Education Endowment Foundation

VERSION

VoR (Version of Record)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Hodgen, Jeremy, Colin Foster, Rachel Marks, and Margaret Brown. 2019. "Evidence for Review of Mathematics Teaching: Improving Mathematics in Key Stages Two and Three". figshare.
<https://hdl.handle.net/2134/36958>.

Improving Mathematics in Key Stages Two and Three: Evidence Review

March 2018

Jeremy Hodgen (UCL Institute of Education)
Colin Foster (University of Leicester)
Rachel Marks (University of Brighton)
Margaret Brown (King's College London)

Please cite this report as follows: Hodgen, J., Foster, C., Marks, R., & Brown, M. (2018). *Evidence for Review of Mathematics Teaching: Improving Mathematics in Key Stages Two and Three: Evidence Review*. London: Education Endowment Foundation. The report is available from:

<https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/improving-mathematics-in-key-stages-two-and-three/>

Contents

1	Introduction	4
2	Executive Summary	7
3	Overview of the development of mathematics competency	16
	3.1 <i>Knowing and learning mathematics</i>	16
	3.1.1 Facts, procedures and concepts	16
	3.1.2 Generic mathematical skills	16
	3.1.3 Building on learners' existing knowledge	17
	3.1.4 Metacognition	17
	3.1.5 Productive dispositions and attitudes	17
	3.2 <i>Teaching and the process of learning</i>	18
	3.2.1 Manipulatives and representations	19
	3.2.2 Teaching strategies	19
	3.2.3 Insights from cognitive science	19
	3.3 <i>Learning trajectories</i>	20
	3.3.1 Variation among learners	20
	3.3.2 Planning for progression.....	20
	3.3.3 Learning trajectories for use in England	21
	3.4 <i>Understanding learners' difficulties</i>	21
	3.4.1 Formative assessment and misconceptions	21
	3.4.2 Developing mathematical competency	22
4	Guide to Reading the Modules	27
	4.1 <i>Meta-analysis, effect sizes and systematic reviews</i>	27
	4.2 <i>Structure</i>	27
5	Method	29
	5.1 <i>Our approach to analysing and synthesising the literature</i>	29
	5.2 <i>Limitations</i>	29
	5.3 <i>Data set</i>	29
	5.4 <i>Coding and data extraction</i>	29
6	Pedagogic Approaches	31
	6.1 <i>Feedback and formative assessment</i>	31
	6.2 <i>Collaborative learning</i>	38
	6.3 <i>Discussion</i>	44
	6.4 <i>Explicit teaching and direct instruction</i>	48
	6.5 <i>Mastery learning</i>	55
	6.6 <i>Problem solving</i>	60
	6.7 <i>Peer and cross-age tutoring</i>	70
	6.8 <i>Misconceptions</i>	77
	6.9 <i>Thinking skills, metacognition and self-regulation</i>	79
7	Resources and Tools	88
	7.1 <i>Calculators</i>	88
	7.2 <i>Technology: technological tools and computer-assisted instruction</i>	93
	7.3 <i>Concrete manipulatives and other representations</i>	101
	7.4 <i>Tasks</i>	108
	7.5 <i>Textbooks</i>	112

8	Mathematical Topics	120
8.1	Overview	120
8.2	Algebra	121
8.3	Number and calculation	128
8.4	Geometry	136
8.5	Probability and Statistics	139
9	Wider School-Level Strategies	141
9.1	Grouping by attainment or 'ability'	141
9.2	Homework	149
9.3	Parental engagement	154
10	Attitudes and Dispositions	160
11	Transition from Primary to Secondary	163
12	Teacher Knowledge and Professional Development	166
13	References	172
14	Appendix: Technical	192
15	Appendix: Literature Searches	193
16	Appendix: Inclusion/Exclusion Criteria	200

1 Introduction

This document presents a review of evidence commissioned by the Education Endowment Foundation to inform the guidance document *Improving Mathematics in Key Stages Two and Three* (Education Endowment Foundation, 2017).

The review draws on a substantial parallel study by the same research team, funded by the Nuffield Foundation, which focuses on the problems faced by low attaining Key Stage three students in developing their maths understanding, and the effectiveness of teaching approaches in overcoming these difficulties. This project, *Low attainment in mathematics: an investigation focusing on Year 9 pupils* includes a systematic review of the evidence relating to teaching of low-attaining secondary students, which the current report builds upon in the wider context of teaching maths in Key Stages two and three.

The Education Endowment Foundation and the Nuffield Foundation are both committed to finding ways of synthesising high quality research about effective teaching and learning, and providing this to practitioners in accessible forms.

There have been a number of recent narrative and systematic reviews of mathematics education examining how students learn and the implications for teaching (e.g., Anthony & Walshaw, 2009; Conway, 2005; Kilpatrick et al., 2001; Nunes et al., 2010). Although this review builds on these studies, this review has a different purpose and takes a different methodological approach to reviewing and synthesising the literature.

The purpose of the review is to synthesise the best available international evidence regarding teaching mathematics to children between the ages of 9 and 14 and to address the question: what is the evidence regarding the effectiveness of different strategies for teaching mathematics?

In addition to this broad research question, we were asked to address a set of more detailed topics developed by a group of teachers and related to aspects of pupil learning, pedagogy, the use of resources, the teaching of specific mathematical content, and pupil attitudes and motivation. Using these topics, we derived the 24 research questions that we address in this review.

Our aim was to focus primarily on robust, causal evidence of impact, using experimental and quasi-experimental designs. However, there are a very large number of experimental studies relevant to this research question. Hence, rather than identifying and synthesising all these primary studies, we focused instead on working with existing meta-analyses and systematic reviews. This approach has the advantage that we can draw on the findings of a very extensive set of original studies that have already been screened for research quality and undergone some synthesis.

Using a systematic literature search strategy, we identified 66 relevant meta-analyses, which synthesise the findings of more than 3000 original studies. However, whilst this corpus of literature is very extensive, there were nevertheless significant gaps. For example, the evidence concerning the teaching of specific mathematical content and topics was limited. In order to address gaps in the meta-analytic

literature, we supplemented our main dataset with 22 systematic reviews identified through the same systematic search strategy.

The structure of this document

We begin with an executive summary with our headline findings. Then, in order to contextualise the review of evidence, we outline our theoretical understanding of how children learn and develop mathematically in Section 3: the development of mathematics competency. In this section, we summarise a range of background literature that we used to inform our analysis and synthesis of the literature.

In Sections 4 and 5, we provide a guide for the reader and describe our method.

In the subsequent sections, we present the findings relating to the 24 detailed research questions. These are organised using a modular approach (as described in Section 4).

Acknowledgements

This review was commissioned by the Education Endowment Foundation (EEF). The EEF is an independent charity dedicated to breaking the link between family income and educational achievement. It generates new evidence about the most effective ways to support disadvantaged pupils; creates free, independent and evidence-based resources; and supports teachers to apply research evidence to their practice. Together with the Sutton Trust, the EEF has been designated the Government's What Works Centre for Educational Outcomes. More information is available at www.educationendowmentfoundation.org.uk.

As explained above, this report builds on a substantial parallel study, examining the evidence relating to the teaching of low-attaining secondary students, funded by the Nuffield Foundation. As such, we were able to draw on a substantial dataset of meta-analyses that we had already identified and summarised for the low attainers study, and we are grateful to the Nuffield Foundation for this funding. Findings from this project will be published later in 2018 and more information about the Nuffield-funded project can be found at <http://www.nuffieldfoundation.org/low-attainment-mathematics-investigation-year-9-students>.

The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project, but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org.

We would like to thank the many peer reviewers who commented on various parts of the review, including Ann Dowker, Sue Gifford, Jenni Golding, Steve Higgins, Ian Jones, Tim Rowland, Ken Ruthven and Anne Watson. We would also like to thank our collaborators on the Nuffield Foundation funded low attainers study, Rob Coe and Steve Higgins, together with the members of the advisory group, Nicola Bretscher, Ann Dowker, Peter Gates, Matthew Inglis, Jane Jones, Anne Watson, John Westwell and Anne White. March 2018.

Reference

Hodgen, J., Foster, C., Marks, R., Brown, M. (2017). Improving Mathematics in Key Stages Two and Three: Evidence Review. London: Education Endowment Foundation. Available at:
<https://educationendowmentfoundation.org.uk/tools/guidance-reports/maths-ks-two-three/>

2 Executive Summary

Feedback and formative assessment (Section 6.1)

What is the effect of giving feedback to learners in mathematics?

The general findings in the EEF toolkit on feedback appear to apply to mathematics: research tends to show that feedback has a large effect on learning, but the range of effects is wide and a proportion of studies show negative effects. The effect of formative assessment is more modest, but is more effective when teachers receive professional development or feedback is delivered through computer-assisted instruction. In mathematics, it may be particularly important to focus on the aspects of formative assessment that involve feedback. Feedback should be used sparingly and predominantly reserved for more complex tasks, where it may support learners' perseverance. The well-established literature on misconceptions and learners' understandings in mathematics provides a fruitful framework to guide assessment and feedback in mathematics. (See 6.8 below.)

Strength of evidence: HIGH

Collaborative learning (Section 6.2)

What is the evidence regarding the effect of using collaborative learning approaches in the teaching and learning of maths?

Collaborative Learning (CL) has a positive effect on attainment and attitude for all students, although the effects are larger at secondary. The largest and most consistent gains have been shown by replicable structured programmes lasting 12 weeks or more. Unfortunately, these programmes are designed for the US educational system, and translating the programmes (and the effects) for the English educational system is not straightforward. The evidence suggests that students need to be taught how to collaborate, and that this may take time and involve changes to the classroom culture. Some English-based guidance is available.

Strength of evidence: HIGH

Discussion (Section 6.3)

What is known about the effective use of discussion in teaching and learning mathematics?

Discussion is a key element of mathematics teaching and learning. However, there is limited evidence concerning the effectiveness of different approaches to improving the quality of discussion in mathematics classrooms. The available evidence suggests that teachers need to structure and orchestrate discussion, scaffold learners' contributions, and develop their own listening skills. Wait time, used appropriately, is an effective way of increasing the quality of learners' talk. Teachers need to emphasise learners' explanations in discussion and support the development of their learners' listening skills.

Strength of evidence: LOW

Explicit teaching and direct instruction (Section 6.4)

What is the evidence regarding explicit teaching as a way of improving pupils' learning of mathematics?

Explicit instruction encompasses a wide array of teacher-led strategies, including direct instruction (DI). There is evidence that structured teacher-led approaches can raise mathematics attainment by a sizeable amount. DI may be particularly beneficial for students with learning difficulties in mathematics. But the picture is complicated, and not all of these interventions are effective. Furthermore, these structured DI programmes are designed for the US and may not translate easily to the English context. Whatever the benefits of explicit instruction, it is unlikely that explicit instruction is effective for all students across all mathematics topics at all times. How the teacher uses explicit instruction is critical, and although careful use is likely to be beneficial, research does not tell us how to balance explicit instruction with other more implicit teaching strategies and independent work by students.

Strength of evidence: MEDIUM

Mastery learning (Section 6.5)

What is the evidence regarding mastery learning in mathematics?

Evidence from US studies in the 1980s generally shows mastery approaches to be effective, particularly for mathematics attainment. However, very small effects were obtained when excluding all but the most rigorous studies carried out over longer time periods. Effects tend to be higher for primary rather than secondary learners and when programmes are teacher-paced, rather than student-paced. The US meta-analyses are focused on two structured mastery programmes, which are somewhat different from the kinds of mastery approaches currently being promoted in England. Only limited evidence is available on the latter, which suggests that, at best, the effects are small. There is a need for more research here.

Strength of evidence: MEDIUM

Problem solving (Section 6.6)

What is the evidence regarding problem solving, inquiry-based learning and related approaches in mathematics?

Inquiry-based learning (IBL) and similar approaches involve posing mathematical problems for learners to solve without teaching a solution method beforehand. Guided discovery can be more enjoyable and memorable than merely being told, and IBL has the potential to enable learners to develop generic mathematical skills, which are important for life and the workplace. However, mathematical exploration can exert a heavy cognitive load, which may interfere with efficient learning. Teachers need to scaffold learning and employ other approaches alongside IBL, including explicit teaching. Problem solving should be an integral part of the

mathematics curriculum, and is appropriate for learners at all levels of attainment. Teachers need to choose problems carefully, and, in addition to more routine tasks, include problems for which learners do not have well-rehearsed, ready-made methods. Learners benefit from using and comparing different problem-solving strategies and methods and from being taught how to use visual representations when problem solving. Teachers should encourage learners to use worked examples to compare and analyse different approaches, and draw learners' attention to the underlying mathematical structure. Learners should be helped to monitor, reflect on and discuss the experience of solving the problem, so that solving the problem does not become an end in itself. At primary school level, it appears to be more important to focus on making sense of representing the problem, rather than on necessarily solving it.

Strength of evidence (IBL): LOW

Strength of evidence (use of problem solving): MEDIUM

Peer and cross-age tutoring (Section 6.7)

What are the effects of using peer and cross-age tutoring on the learning of mathematics?

Peer and cross-age tutoring appear to be beneficial for tutors, tutees and teachers and involve little monetary cost, potentially freeing up the teacher to implement other interventions. Cross-age tutoring returns higher effects, but is based on more limited evidence. Peer-tutoring effects are variable, but are not negative. Caution should be taken when implementing tutoring approaches with learners with learning difficulties.

Strength of evidence: MEDIUM

Misconceptions (Section 6.8)

What is the evidence regarding misconceptions in mathematics?

Students' misconceptions arise naturally over time as a result of their attempts to make sense of their growing mathematical experience. Generally, misconceptions are the result of over-generalisation from within a restricted range of situations. Misconceptions should be viewed positively as evidence of students' sense making. Rather than confronting misconceptions in an attempt to expunge them, exploration and discussion can reveal to students the limits of applicability associated with the misconception, leading to more powerful and extendable conceptions that will aid students' subsequent mathematical development.

Strength of evidence: MEDIUM

Thinking skills, metacognition and self-regulation (Section 6.9)

To what extent does teaching thinking skills, metacognition and/or self-regulation improve mathematics learning?

Teaching thinking skills, metacognition and self-regulation can be effective in mathematics. However, there is a great deal of variation across studies. Implementing these approaches is not straightforward. The development of thinking skills, metacognition and self-regulation takes time (more so than other concepts), the duration of the intervention matters, and the role of the teacher is important. One thinking skills programme developed in England, Cognitive Acceleration in Mathematics Education (CAME), appears to be particularly promising. Strategies that encourage self-explanation and elaboration appear to be beneficial. There is some evidence to suggest that, in primary, focusing on cognitive strategies may be more effective, whereas, in secondary, focusing on learner motivation may be more important. Working memory and other aspects of executive function are associated with mathematical attainment, although there is no clear evidence for a causal relationship. A great deal of research has focused on ways of improving working memory. However, whilst working memory training improves performance on tests of working memory, it does not have an effect on mathematical attainment.

Strength of evidence (Thinking skills, metacognition and self-regulation): MEDIUM

Strength of evidence (Working memory training): HIGH

Calculators (Section 7.1)

What are the effects of using calculators to teach mathematics?

Calculator use does not in general hinder students' skills in arithmetic. When calculators are used as an integral part of testing and teaching, their use appears to have a positive effect on students' calculation skills. Calculator use has a small positive impact on problem solving. The evidence suggests that primary students should not use calculators every day, but secondary students should have more frequent unrestricted access to calculators. As with any strategy, it matters how teachers and students use calculators. When integrated into the teaching of mental and other calculation approaches, calculators can be very effective for developing non-calculator computation skills; students become better at arithmetic in general and are likely to self-regulate their use of calculators, consequently making less (but better) use of them.

Strength of evidence: HIGH

Technology: technological tools and computer-assisted instruction (Section 7.2)

What is the evidence regarding the use of technology in the teaching and learning of maths?

Technology provides powerful tools for representing and teaching mathematical ideas. However, as with tasks and textbooks, how teachers use technology with learners is critical. There is an extensive research base examining the use of computer-assisted instruction (CAI), indicating that CAI does not have a negative effect on learning. However, the research is almost exclusively focused on systems

designed for use in the US in the past, some of which are now obsolete. More research is needed to evaluate the use of CAI in the English context.

Strength of evidence (Tools): LOW

Strength of evidence (CAI): MEDIUM

Concrete manipulatives and other representations (Section 7.3)

What are the effects of using concrete manipulatives and other representations to teach mathematics?

Concrete manipulatives can be a powerful way of enabling learners to engage with mathematical ideas, provided that teachers ensure that learners understand the links between the manipulatives and the mathematical ideas they represent. Whilst learners need extended periods of time to develop their understanding by using manipulatives, using manipulatives for too long can hinder learners' mathematical development. Teachers need to help learners through discussion and explicit teaching to develop more abstract, diagrammatic representations. Number lines are a particularly valuable representational tool for teaching number, calculation and multiplicative reasoning across the age range. Whilst in general the use of multiple representations appears to have a positive impact on attainment, the evidence base concerning specific approaches to teaching and sequencing representations is limited. Comparison and discussion of different representations can help learners develop conceptual understanding. However, using multiple representations can exert a heavy cognitive load, which may hinder learning. More research is needed to inform teachers' choices about which, and how many, representations to use and when.

Strength of evidence (Manipulatives): HIGH

Strength of evidence (Representations): MEDIUM

Tasks (Section 7.4)

What is the evidence regarding the effectiveness of mathematics tasks?

The current state of research on mathematics tasks is more directly applicable to curriculum designers than to schools. Tasks frame, but do not determine, the mathematics that students will engage in, and should be selected to suit the desired learning intentions. However, as with textbooks, how teachers use tasks with students is more important in determining their effectiveness. More research is needed on how to communicate the critical pedagogic features of tasks so as to enable teachers to make best use of them in the classroom.

Strength of evidence: LOW

Textbooks (Section 7.5)

What is the evidence regarding the effectiveness of textbooks?

The effect on student mathematical attainment of using one textbook scheme rather than another is very small, although the choice of a textbook will have an impact on what, when and how mathematics is taught. However, in terms of increasing mathematical attainment, it is more important to focus on professional development and instructional differences rather than on curriculum differences. The organisation of the mathematics classroom and how textbooks can enable teachers to develop students' understanding of, engagement in and motivation for mathematics is of greater significance than the choice of one particular textbook rather than another.

Strength of evidence: HIGH

Algebra (Section 8.2)

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of algebra?

Learners generally find algebra difficult because of its abstract and symbolic nature and because of the underlying structural features, which are difficult to operate with. This is especially the case if learners experience the subject as a collection of arbitrary rules and procedures, which they then misremember or misapply. Learners benefit when attention is given both to procedural and to conceptual teaching approaches, through both explicit teaching and opportunities for problem-based learning. It is particularly helpful to focus on the structure of algebraic representations and, when solving problems, to assist students in choosing deliberately from alternative algebraic strategies. In particular, worked examples can help learners to appreciate algebraic reasoning and different solution approaches.

Strength of evidence: MEDIUM

Number and calculation (Section 8.3)

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of number and calculation?

Number and numeric relations are central to mathematics. Teaching should enable learners to develop a range of mental and other calculation methods. Quick and efficient retrieval of number facts is important to future success in mathematics. Fluent recall of procedures is important, but teaching should also help learners understand how the procedures work and when they are useful. Direct, or explicit, teaching can help learners struggling with number and calculation. Learners should be taught that fractions and decimals are numbers and that they extend the number system beyond whole numbers. Number lines should be used as a central representational tool in teaching number, calculation and multiplicative reasoning across Key Stages 2 and 3.

Strength of evidence: MEDIUM

Geometry (Section 8.4)

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of geometry and measures?

There are few studies that examine the effects of teaching interventions for and pedagogic approaches to the teaching of geometry. However, the research evidence suggests that representations and manipulatives play an important role in the learning of geometry. Teaching should focus on conceptual as well as procedural knowledge of measurement. Learners experience particular difficulties with area, and need to understand the multiplicative relations underlying area.

Strength of evidence: LOW

Probability and Statistics (Section 8.5)

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of probability and statistics?

There are very few studies that examine the effects of teaching interventions for and pedagogic approaches to the teaching of probability and statistics. However, there is research evidence on the difficulties that learners experience and the common misconceptions that they encounter, as well as the ways in which they learn more generally. This evidence suggests some pedagogic principles for the teaching of statistics.

Strength of evidence: LOW

Grouping by attainment or 'ability' (Section 9.1)

What is the evidence regarding 'ability grouping' on the teaching and learning of maths?

Setting or streaming students into different classes for mathematics based on their prior attainment appears to have an overall neutral or slightly negative effect on their future attainment, although higher attainers may benefit slightly. The evidence suggests no difference for mathematics in comparison to other subjects. The use of within-class grouping at primary may have a positive effect, particularly for mathematics, but if used then setting needs to be flexible, with regular opportunities for group reassignment.

Strength of evidence: MEDIUM

Homework (Section 9.2)

What is the evidence regarding the effective use of homework in the teaching and learning of mathematics?

The effect of homework appears to be low at the primary level and stronger at the secondary level, although the evidence base is weak. It seems to matter more that homework encourages students to actively engage in learning rather than simply

learning by rote or finishing off classwork. In addition, the student's effort appears to be more important than the time spent or the quantity of work done. This would suggest that the teacher should aim to set homework that students find engaging and that encourages metacognitive activity. For primary students, homework seems not to be associated with improvements in attainment, but there could be other reasons for setting homework in primary, such as developing study skills or student engagement. Homework is more important for attainment as students get older. As with almost any intervention, teachers make a huge difference. It is likely that student effort will increase if teachers value students' homework and discuss it in class. However, it is not clear that spending an excessive amount of time marking homework is an effective use of teacher time.

Strength of evidence: LOW

Parental engagement (Section 9.3)

What is the evidence regarding parental engagement and learning mathematics?

The well-established association between parental involvement and a child's academic success does not appear to apply to mathematics, and there is limited evidence on how parental involvement in mathematics might be made more effective. Interventions aimed at improving parental involvement in homework do not appear to raise attainment in mathematics, and may have a negative effect in secondary. However, there may be other reasons for encouraging parental involvement. Correlational studies suggest that parental involvement aimed at increasing academic socialization, or helping students see the value of education, may have a positive impact on achievement at secondary.

Strength of evidence: LOW

Attitudes and Dispositions (Section 10)

How can learners' attitudes and dispositions towards mathematics be improved and maths anxiety reduced?

Positive attitudes and dispositions are important to the successful learning of mathematics. However, many learners are not confident in mathematics. There is limited evidence on the efficacy of approaches that might improve learners' attitudes to mathematics or prevent or reduce the more severe problems of maths anxiety. Encouraging a growth mindset rather than a fixed mindset is unlikely to have a negative impact on learning and may have a small positive impact.

Strength of evidence: LOW

Transition from Primary to Secondary (Section 11)

What is the evidence regarding how teaching can support learners in mathematics across the transition between Key Stage 2 and Key Stage 3?

The evidence indicates a large dip in mathematical attainment as children move from primary to secondary school in England, which is accompanied by a dip in learner attitudes. There is very little evidence concerning the effectiveness of particular interventions that specifically address these dips. However, research does indicate that initiatives focused on developing shared understandings of curriculum, teaching and learning are important. Both primary and secondary teachers are likely to be more effective if they are familiar with the mathematics curriculum and teaching methods outside of their age phase. Secondary teachers need to revisit key aspects of the primary mathematics curriculum, but in ways that are engaging and relevant and not simply repetitive. Teachers' beliefs about their ability to teach appear to be particularly crucial for lower-attaining students in Key Stage 3 mathematics.

Strength of evidence: LOW

Teacher Knowledge and Professional Development (Section 12)

What is the evidence regarding the impact of teachers and their effective professional development in mathematics?

The evidence shows that the quality of teaching makes a difference to student outcomes. The quality of teaching, or instructional guidance, is important to the efficacy of almost every strategy that we have examined. The evidence also indicates that, in mathematics, teacher knowledge is a key factor in the quality of teaching. Teacher knowledge, more particularly pedagogic content knowledge (PCK), is crucial in realising the potential of mathematics curriculum resources and interventions to raise attainment. Professional development (PD) is key to raising the quality of teaching and teacher knowledge. However, evidence concerning the specific effects of PD is limited. This evidence suggests that extended PD is more likely to be effective than short courses.

Strength of evidence (Teacher knowledge): LOW

Strength of evidence (Teacher PD): LOW

3 Overview of the development of mathematics competency

In this section, we describe in broad terms how learners typically develop competency in mathematics. We conceptualise ‘typical’ as the common range of developmental trajectories demonstrated by the majority of learners in mainstream primary and secondary education in England. We note that there is a wide variation in learners’ mathematical development and that it is helpful to conceive of this variation as a continuum (Brown, Askew et al., 2008). However, whilst the range in development and attainment is wide, many children experience similar difficulties.

3.1 Knowing and learning mathematics

Successful learning of mathematics requires several elements to be in place, which together enable the learner to make progress, navigate difficulties and develop mathematics competency.

3.1.1 Facts, procedures and concepts

It is helpful to think of mathematical knowledge as consisting of factual, procedural and conceptual knowledge, which are strongly inter-related (Donovan & Bransford, 2005; Kilpatrick et al., 2001). To become mathematically competent, learners need to develop a rich foundation of factual and procedural knowledge. However, while knowing how to carry out a procedure fluently is important, it is not sufficient; learners also need to identify *when* the procedure is appropriate, understand *why* it works and know *how* to interpret the result (Hart et al., 1981). This requires conceptual knowledge,¹ which involves understanding the connections and relationships between mathematical facts, procedures and concepts; for example, understanding addition and subtraction as inverse operations (Nunes et al., 2009). Additionally, learners need to organise their knowledge of facts, procedures and concepts in ways that enable them to retrieve and apply this knowledge, although we emphasise that this organisation is largely unconscious.² Nunes et al. (2012) refer to the use of conceptual knowledge as *mathematical reasoning*, and have shown this to be an important predictor of future mathematical attainment. Similarly, Dowker (2014) has demonstrated a strong relationship between calculational proficiency and the extent to which children use *derived fact* strategies based on conceptual links (e.g., if $67 - 45 = 22$, then $68 - 45$ must be 23), whilst Gray & Tall (1994) found that higher-attaining students used strategies such as these as part of their progression towards competent calculation, whereas lower-attaining students did not.

The relations between how factual, procedural and conceptual knowledge are learnt, however, are contested. For example, in devising curriculum sequences it is often assumed that conceptual knowledge should be placed before the associated procedural knowledge, so that the concepts can support the procedures (NCTM, 2014), but there is evidence that procedural knowledge can also support conceptual knowledge, and therefore that these kinds of knowledge are mutually interdependent (Rittle-Johnson, Schneider, & Star, 2015).

3.1.2 Generic mathematical skills

To solve problems, learners need to develop generic mathematical strategies, sometimes known as ‘processes’ or ‘generic mathematical skills’ (HMI, 1985), or as

‘strategic competence’, which Kilpatrick et al. (2001) define as the “ability to formulate, represent, and solve mathematical problems” (p. 5). These include actions such as specialising and generalising, and conjecturing and proving (Mason & Johnston-Wilder, 2006, pp. 74-77). The development of these strategies appears to be supported by teachers highlighting when they or their learners spontaneously use them; for example, by naming them and asking for other examples of their use (Mason, 2008).

3.1.3 Building on learners’ existing knowledge

Learners come to mathematics classrooms with existing mathematical knowledge and preconceptions, much of which is useful and at least partially effective. In order to develop mathematics competence, teaching needs to enable learners to build upon, transform and restructure their existing knowledge (Donovan & Bransford, 2005; see also Bransford et al., 2000). This is particularly important where such preconceptions, or ‘metbefore’s’ (McGowen & Tall, 2010), are likely to interfere with learning (see Section 4.1 below).

3.1.4 Metacognition

Learning and doing mathematics involves more than knowledge and cognitive activity. Fostering metacognition appears to be important to the development of mathematics competence (Donovan & Bransford, 2005). Metacognition is defined in different ways by different researchers (Gascoine et al., 2017), some focusing on “thinking about thinking” (Adey & Shayer, 1994) and others on “learning to learn” (see discussion in Higgins et al., 2005). Donovan & Bransford (2005) define metacognition as “the phenomenon of ongoing sense making, reflection, and explanation to oneself and others” (p. 218) and equate it to Kilpatrick et al.’s (2001) “adaptive reasoning [which is] ... the capacity to think logically about the relationships among concepts and situations and to justify and ultimately prove the correctness of a mathematical procedure or assertion ... [which] includes reasoning based on pattern, analogy or metaphor” (p. 170). Mathematics-specific metacognitive activity is distinct from generic metacognitive approaches. Metacognition related to mathematics includes a generic component (logical thinking, including induction, deduction, generalisation, specialisation, etc.) as well as a mathematics-specific component (e.g., identifying relationships between variables and expressing them in tables, graphs and symbols). Mathematical discussion and dialogue can support metacognitive activity (Donovan & Bransford, 2005),³ and mathematical discussion is more than just talk. Learners benefit from being taught how to engage in discussion (Kyriacou & Issitt, 2008), and orchestrating productive mathematical discussions requires considerable pedagogical skill (Stein et al., 2008).

3.1.5 Productive dispositions and attitudes

Successful learning also depends on learners’ attitudes and productive dispositions towards mathematics, as well as contributing to these. Attitudes can be defined as “a liking or disliking of mathematics, a tendency to engage in or avoid mathematical activities, a belief that one is good or bad at mathematics, and a belief that mathematics is useful or useless” (Neale, cited in Ma & Kishnor, 1997, p.27). The relationship between learners’ attitudes and attainment is weak but important, and

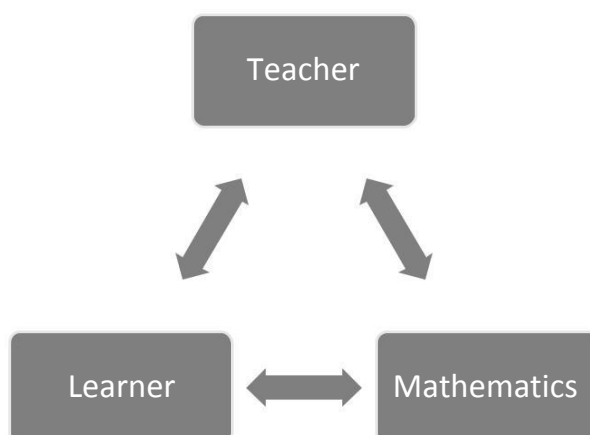
attitudes become increasingly negative as learners get older (Ma & Kishnor, 1997). Attitudes appear to be an important factor in progression and participation in mathematics post-16 (Brown, Brown & Bibby, 2008). Some learners experience maths anxiety, which can be a very strong hindrance to learning and doing mathematics (Dowker et al., 2016; see also Chinn, 2009). Estimates of the extent of maths anxiety vary considerably from 2-6% among secondary-school pupils in England (Chin, 2009) to 68% of US college students registered on mathematics courses (Betz, cited in Dowker et al., 2016).

Kilpatrick et al. (2001) describe *productive dispositions* as the “habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one’s own efficacy” (p. 5) and, thus, as encompassing more than attitudes. These include motivation (Middleton & Spanias, 1999), mathematical resilience (Johnston-Wilder & Lee, 2010), mathematical self-efficacy, the belief in one’s ability to carry out an activity (Bandura & Schunk, 1981) as well as beliefs about the value of mathematics. Productive mathematical activity requires self-regulation, which, for the purposes of this review, is defined as the dispositions required to control one’s emotions, thinking and behaviour, including one’s cognitive and metacognitive actions (Dignath & Büttner, 2008; see also Gascoine et al., 2016).⁴

Emerging research suggests the importance of particular dispositions towards mathematics, such as ‘spontaneous focusing’ on number, mathematical relations or patterns, although it is not clear how, and to what extent, such dispositions are amenable to teaching (e.g., Rathé et al., 2016; Verschaffel, forthcoming).

3.2 Teaching and the process of learning

The teacher, the learner and the mathematics can be conceptualised as a dynamic *teaching triad* (see figure, based on Steinbring, 2011, p. 44), in which the teacher mediates between the learner and the mathematics by providing tasks, resources and representations to help the learner to make sense of the mathematics. (The teaching triad, or ‘didactical triangle’ was originally suggested by Herbart, see Steinbring, 2011.)



3.2.1 Manipulatives and representations

Manipulatives (concrete materials) and other representations offer powerful support for learners, which may be gradually internalised as mental images take over (Streefland, 1991; Carbonneau et al., 2013). However, teachers need to help learners to link the materials (and the actions performed on or with them) to the mathematics of the situation, to appreciate the limitations of concrete materials, and to develop related mathematical images, representations and symbols (Nunes et al., 2009). In a similar way, diagrams and models that enable learners to build on their intuitive understandings of situations can be powerful ways of approaching mathematical problems (Nunes et al., 2009). Such models of problems can then become more powerful models *for* understanding and tackling problems with related mathematical structure, where it may be less straightforward to use one's intuition (Streefland, 1991; see also Nunes et al., 2009). But, while time and experience are necessary elements for this process to occur, learners cannot be left entirely to 'discover' these powerful models for themselves; transforming intuitive representations in this way requires some explicit teaching and structured discussion (e.g., see Askew et al., 1997; Kirschner et al., 2006).

3.2.2 Teaching strategies

It seems likely that the effectiveness of different teaching strategies will depend on the particular aspects of mathematical knowledge in question, as well as on learner differences. For example, explicit/direct instruction (Gersten, Woodward, & Darch, 1986) could be particularly effective for teaching particular procedures at particular points in learners' mathematical development, but might be less effective at developing reasoning, addressing persistent misconceptions or supporting metacognition. In Part 2 of this review of teaching strategies, we examine evidence for the relative efficacy of different strategies relating to different aspects of mathematics competence.

3.2.3 Insights from cognitive science

There is currently a great deal of interest in how insights from cognitive science (e.g. cognitive load theory) may be relevant to mathematics teaching and learning (Alcock et al., 2016; Gilmore et al., forthcoming; Wiliam, 2017). *Cognitive load theory* (CLT) originated in the 1980s and addresses the instructional implications of the demands that are placed on *working memory* (Sweller, 1994). All conscious cognitive processing takes place in working memory, which is highly limited and able to handle only a small number of novel interacting elements at a time – far fewer than the number normally needed for most kinds of sophisticated intellectual activity. In contrast, *long-term memory* allows us to store an almost limitless number of *schemas*, which are cognitive constructs that chunk multiple pieces of information into a single element (Paas, Renkl, & Sweller, 2003). When a schema is brought from long-term memory into working memory, even though it consists of a complex set of interacting elements, it can be processed as just *one* element. In this way, far more sophisticated processing can take place than would be possible with working memory alone. Important findings include the *expertise reversal effect*, in which “instructional techniques that are effective with novices can lose their effectiveness and even become ineffective when used with more experienced learners” (Paas, Renkl, & Sweller, 2003, p. 3), the *worked examples effect*, in which cognitive load is

reduced by studying worked, or partially worked, examples rather than solving the equivalent problems, and the *generation effect*, whereby learners better remember ideas that they have at least partially created for themselves (Chen et al., 2015).

3.3 Learning trajectories

While there is considerable variation in what different children learn when, there are some overall trends in children's learning, which are captured by the notion of learning trajectories (Clements & Sarama, 2004). Learning trajectories (or learning progressions) are "empirically supported hypotheses about the levels or waypoints of thinking, knowledge, and skill in using knowledge, that [learners] are likely to go through as they learn mathematics" (Daro et al., 2011, p. 12).

3.3.1 Variation among learners

Learners vary considerably in their levels of attainment and understanding. Children differ in how long it takes them to come to know mathematics; e.g. the gap in typical attainment⁵ is equivalent to approximately 7-8 years' learning by the time learners reach Key Stage 3 (Cockcroft, 1982; see also Brown, Askew et al., 2008; Jerrim & Shure, 2016). There can clearly be no expectation that all learners will progress through the key waypoints at the same time, or even necessarily in the same order. Learning can appear idiosyncratic and non-linear, with learners at any one time sometimes more likely to succeed with an apparently more complex idea than with a simpler one. Difficult ideas may initially be learned at a superficial level and must then be returned to, perhaps many times, before deep conceptual understanding develops and is retained (Denvir & Brown, 1986; Brown et al., 1995; Pirie and Kieren, 1994). Classroom learning is the product of interactions between teachers, learners and mathematics (Kilpatrick et al., 2001) and is dependent on learners' prior experiences, interests and motivations. Indeed, differences in the taught curriculum, home and society between, for example, England and the US, or the Pacific Rim, are important when considering research evidence from different parts of the world. The possibility of curriculum and other cultural effects must always be borne in mind, and findings cannot be transplanted simplistically from one place to another (Askew et al., 2010).

3.3.2 Planning for progression

Consequently, no single learning trajectory can describe the development of all learners at all times. However, there are broad patterns of progression in many skills. For example, when learning about addition, we would expect the vast majority of learners to *count all* before moving to *count on* (e.g., Gravemeijer, 1994). As we have already observed, it is helpful to consider most children's mathematical development as falling on a continuum of typical development. Effective planning of a curriculum, as well as effective planning of support for all learners, needs to engage with realistic expectations regarding the likely variation in learning trajectories, and to encourage the development of strategies at different levels. Key to this is the way in which teachers themselves conceive of, and teach, mathematics as a connected discipline (Askew et al., 1997; Hiebert & Carpenter, 1992).

3.3.3 Learning trajectories for use in England

There are several research-based approaches to learning trajectories developed in England (e.g. Brown, 1992), the US (e.g. Clements & Samara, 2014; Confrey et al., 2009) and elsewhere (e.g., Clarke et al., 2000; de Lange, 1999). Some focus on particular strands or stages, such as primary number (Clarke et al., 2000) or multiplicative reasoning (Confrey et al., 2009). Learning trajectories have been comprehensively described in English policy documents, such as various versions of the National Curriculum (Brown, 1996) and the Primary and Secondary Frameworks for Teaching Mathematics (DfEE, 1998; 2001). However, it is important to note that none of these documents is perfect and, as Daro, Mosher and Corcoran (2011) observe, “There are major gaps in our understanding of learning trajectories in mathematics” (p. 13). The learning trajectories described in English policy documents (DfEE, 1998; 2001) provide a model that, despite recent changes to the curriculum, is applicable, with adaptation, to the current English context, and which is at least partially evidence-based (Brown 1989, 1996; Brown et al., 1998). However, these should be read in conjunction with research-based commentaries on teaching and learning, such as Hart et al. (1981), Nunes et al. (2009), Ryan & Williams (2007) and Watson et al. (2013).

3.4 Understanding learners’ difficulties

It is essential for practitioners to understand the different ways in which learners’ mathematics may develop. We will outline models that seem to be most beneficial in allowing practitioners to identify key areas where learners encounter difficulties, as well as effective strategies for addressing these. Formative assessment entails establishing students’ difficulties and adapting teaching so as to respond effectively (Black & Wiliam, 2009).

3.4.1 Formative assessment and misconceptions

We have highlighted the need to understand and build on learners’ existing knowledge. Assessing this knowledge involves being attuned to what learners bring to the mathematics classroom, being able to actively listen to and respond to learners’ own informal strategies (Carpenter et al., 1999) and to have awareness of the mathematical knowledge that learners develop in their everyday lives, such as informal ‘sharing’ (division) practices (Nunes & Bryant, 2009). As part of this, practitioners need knowledge of common errors and misconceptions in mathematics, which are invaluable in diagnosing the difficulties learners encounter (Dickson et al., 1984; Hart, 1981; Ryan & Williams, 2007).

It is important to note that ‘misconceptions’ is a contested term (Daro et al., 2011; Smith III et al., 1994). For the purposes of this review, we define misconceptions as the result of an attempt to make sense of a situation, using ideas that have worked in past situations but do not adequately fit the current one. Hence, the term encompasses various ‘met-befores’ (McGowen & Tall, 2010), such as partial understandings, over-generalisations and incorrect reasoning. It is important for practitioners to recognise misconceptions as part of typical mathematical development, and not necessarily as things that must be avoided or ‘fixed’ immediately. For example, it would be hard to envisage a typical development that

did not include multiplication-makes-bigger-division-makes-smaller at some point along the way (Greer, 1994).

3.4.2 Developing mathematical competency

Each learner's trajectory through mathematics will be to some extent unique, involving their own particular difficulties and successes. However, there are many features of developing competency in mathematics that are common across a wide range of learners. Familiarity with some of the broad findings from research, as summarised in this report, can assist teachers in leading learners confidently through their mathematical journeys and responding in sensitive and mathematically coherent ways when difficulties arise.

References

- Adey, P. S., & Shayer, M. (1994). *Really raising standards*. London: Routledge.
- Alcock, L., Ansari, D., Batchelor, S., Bisson, M.-J., De Smedt, B., Gilmore, C., . . . Weber, K. (2016). Challenges in mathematical cognition: a collaboratively derived research agenda. *Challenges in Mathematical Cognition*, 1(2), 20-41.
- Askew, M., Brown, M., Rhodes, V., Johnson, D. C., & Wiliam, D. (1997). *Effective teachers of numeracy*. London: King's College.
- Askew, M., Hodgen, J., Hossain, S., & Bretscher, N. (2010). *Values and variables: A review of mathematics education in high-performing countries*. London: The Nuffield Foundation.
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of personality and social psychology*, 41(3), 586.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience, and School* (Expanded Edition). Washington, DC: National Academies Press.
- Brown, M. (1996). The context of the research: the evolution of the national curriculum for mathematics. In D. C. Johnson & A. Millett (Eds.), *Implementing the Mathematics National Curriculum: policy, politics and practice* (pp. 1-28). London: Paul Chapman Publishing.
- Brown, M. (Ed.) (1992). *Graded Assessment in Mathematics (GAIM) topic criteria handbook*. Walton on Thames: Nelson.
- Brown, M. (1989). Graded Assessment and Learning Hierarchies in Mathematics: An alternative view. *British Educational Research Journal*, 15(2), 121 - 128. doi:10.1080/0141192890150202
- Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385.
- Brown, M., Askew, M., Hodgen, J., Rhodes, V., Millett, A., Denvir, H., & Wiliam, D. (2008). Individual and cohort progression in learning numeracy ages 5-11: Results from the Leverhulme 5-year longitudinal study. In A. Dowker (Ed.), *Mathematical Difficulties: Psychology and Intervention* (pp. 85-108). Oxford: Elsevier.
- Brown, M., Brown, P., & Bibby, T. (2008). "I would rather die": Attitudes of 16 year-

- olds towards their future participation in mathematics. *Research in Mathematics Education*, 10(1), 3-18.
- Brown, M., Blondel, E., Simon, S., & Black, P. (1995). Progression in measuring. *Research Papers in Education*, 10(2), 143-170. doi:10.1080/0267152950100202
- 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(2), 380. doi: 10.1037/a0031084
- Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann.
- Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology*, 107(3), 689.
- Chinn, S. (2009). Mathematics anxiety in secondary students in England. *Dyslexia*, 15(1), 61-68.
- Clarke, D., Sullivan, P., Cheeseman, J., & Clarke, B. (2000). The Early Numeracy Research Project: Developing a framework for describing early numeracy learning. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000 (Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia)* (pp. 180-187). Freemantle, Western Australia: MERGA.
- Clements, D. H., & Samara, J. (2014). *Learning and Teaching Early Math: The Learning Trajectories Approach*. Oxford: Routledge.
- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical thinking and learning*, 6(2), 81-89.
- Cockcroft, W. H. (1982). *Mathematics counts*. London: HMSO.
- Confrey, J., Maloney, A. P., Nguyen, K. H., Mojica, G., & Myers, M. (2009). Equipartitioning / splitting as a foundation of rational number reasoning using learning trajectories. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 345-352). Thessaloniki, Greece: PME.
- Daro, P., Mosher, F. A., & Corcoran, T. (2011). *Learning trajectories in mathematics: a foundation for standards, curriculum, assessment, and instruction*. Philadelphia, PA: Consortium for Policy Research in Education (CPRE).
- de Lange, J. (1999). *Framework for classroom assessment in mathematics*. Utrecht: Freudenthal Institute & National Center for Improving Student Learning and Achievement in Mathematics and Science.
- Denvir, B., & Brown, M. (1986). Understanding number concepts in low attaining 7-9 year olds. Part II: The teaching studies. *Educational Studies in Mathematics*, 17, 143-164.
- DfEE. (1998). *Framework for teaching mathematics: Reception to Year 6* (September ed.). London: Department for Education and Employment.
- DfEE. (2001). *Framework for teaching mathematics Years 7, 8 and 9*. London: Department for Education and Employment.
- Dickson, L., Brown, M., & Gibson, O. (1984). *Children learning mathematics: A teacher's guide to recent research*. Eastbourne: Holt, Rinehart & Winston

- Ltd.
- Donovan, M. S., & Bransford, J. D. (2005). *How students learn: Mathematics in the classroom*. Washington, DC: The National Academies Press.
- Dowker, A. (2014). Young children's use of derived fact strategies in addition and subtraction. *Frontiers in Human Neuroscience*, 7. doi:10.3389/fnhum.2013.00924
- Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics Anxiety: What Have We Learned in 60 Years? *Frontiers in Psychology*, 7(508). doi:10.3389/fpsyg.2016.00508
- Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, 5(1), 3-57. doi:10.1002/rev3.3077
- Gersten, R., Woodward, J., & Darch, C. (1986). Direct instruction: A research-based approach to curriculum design and teaching. *Exceptional Children*, 53(1), 17-31.
- Gilmore, C.K., Keeble, S., Richardson, S., Cragg, L. (Forthcoming) The Interaction of Procedural Skill, Conceptual Understanding and Working Memory in Early Mathematics Achievement, *Journal of Numerical Cognition*.
- Gravemeijer, K. P. E. (1994). *Developing Realistic Mathematics Education*. Utrecht: CD-β Press / Freudenthal Institute.
- Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for research in Mathematics Education*, 116-140.
- Greer, B. (1994). Extending the Meaning of Multiplication and Division. In G. Harel & J. Confrey (Eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics* (pp. 61-85): SUNY Press.
- Hart, K., Brown, M. L., Küchemann, D. E., Kerslake, D., Ruddock, G., & McCartney, M. (Eds.). (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- HMI. (1985). *Mathematics from 5 to 16*. HMI Series: Curriculum Matters No. 3. London: HMSO.
- Jaworski, B. (1994). *Investigating mathematics teaching: A constructivist enquiry*. London: Falmer Press.
- Jerrim, J., & Shure, N. (2016). *Achievement of 15-Year-Olds in England: PISA 2015 National Report*. London: Department for Education.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding It Up: Helping Children Learn Mathematics (Prepared by the Mathematics Learning Study Committee, National Research Council)*. Washington DC: The National Academies Press.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist,

- Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), 75-86. doi:10.1207/s15326985ep4102_1
- Kyriacou, C., & Issitt, J. (2008). *What Characterises Effective Teacher-initiated Teacher-pupil Dialogue to Promote Conceptual Understanding in Mathematics Lessons in England in Key Stages 2 and 3: A Systematic Review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: a meta-analysis. *Journal for Research In Mathematics Education*, 28(1), 26-47.
- Mason, J. (2008). Being mathematical with and in front of learners: Attention, awareness, and attitude as sources of difference between teacher educators, teachers and learners. In B. Jaworski (Vol. Ed.) & T. Wood (Series Ed.), *Handbook of mathematics teacher education (Vol. 4): The mathematics teacher educator as a developing professional* (pp. 31–56). Rotterdam, Netherlands: Sense.
- Mason, J., & Johnston-Wilder, S. (2006). *Designing and using mathematical tasks*. Tarquin Pubns.
- McGowen, M. A., & Tall, D. O. (2010). Metaphor or Met-Before? The effects of previous experience on practice and theory of learning mathematics. *The Journal of Mathematical Behavior*, 29(3), 169-179.
- Middleton, J. A., & Spanias, P. A. (1999). Motivation for achievement in mathematics: Findings, generalizations and criticisms of the research. *Journal for Research in Mathematics Education*, 30(1), 65-88.
- Nunes, T., & Bryant, P. (2009). Paper 3: Understanding rational numbers and intensive quantities. In T. Nunes, P. Bryant, & A. Watson (Eds.), *Key understandings in mathematics learning*. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 9 October 2009.
- Nunes, T., Bryant, P., Barros, R., & Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. *British Journal of Educational Psychology*, 82(1), 136-156. doi:10.1111/j.2044-8279.2011.02033.x
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1), 1-4.
- Pirie, S. E. B., & Kieren, T. E. (1994). Growth in mathematical understanding: How can we characterise it and how can we represent it? *Educational Studies in Mathematics*, 26, 165 - 190.
- Rathé, S., Torbeyns, J., Hannula-Sormunen, M., De Smedt, B., & Verschaffe, L. (2016). Spontaneous Focusing On Numerosity: A Review of Recent Research. *Mediterranean Journal for Research in Mathematics Education*.
- Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, 27(4), 587-597.

- Ryan, J., & Williams, J. (2007). *Children's mathematics 4-15: learning from errors and misconceptions: learning from errors and misconceptions*. McGraw-Hill Education (UK).
- Smith III, J. P., Disessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The journal of the learning sciences*, 3(2), 115-163.
- Star, J. R. (2005). Research Commentary: Reconceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education*, 36(5), 404-411.
- Stein, M., Engle, R., Smith, M., & Hughes, E. (2008). Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell. *Mathematical Thinking And Learning*, 10(4), 313-340.
- Steinbring, H. (2011). Changed views on mathematical knowledge in the course of didactical theory development: independent corpus of scientific knowledge or result of social constructions? In T. Rowland & K. Ruthven (Eds.), *Mathematical Knowledge in Teaching* (pp. 43-64). Dordrecht, NL: Springer.
- Streefland, L. (1991). *Fractions in realistic mathematics education: A paradigm of developmental research* (Vol. 8). Dordrecht. The Netherlands: Kluwer Academic Publishers.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4), 295-312.
- Verschaffel, L., Torbeyns, J., & De Smedt, B. (Forthcoming). Young children's early mathematical competencies: analysis and stimulation. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10)*. Dublin: European Society for Research in Mathematics Education.
- Watson, A., Jones, K., & Pratt, D. (2013). *Key Ideas in Teaching Mathematics: Research-based guidance for ages 9-19*. Oxford: Oxford University Press.
- William, D. (2017). Memories are made of this. *TES* (2nd June 2017).

Note

¹ Conceptual knowledge, or understanding, is referred to in different ways by different researchers. Kilpatrick et al. (2001) define it as "comprehension of mathematical concepts, operations, and relations" (p. 5). Skemp (1976) refers to relational understanding and distinguishes this from instrumental knowledge, whereas Gray and Tall (1994) focus on a 'procept' as a process/object amalgam. Ma (1999) refers to a profound understanding of fundamental mathematics (although her work is focused on teacher knowledge). Nunes et al. (2012, see also Nunes et al., 2009) refer to mathematical reasoning, while others distinguish deep from superficial knowledge (Star, 2005). Hart et al. (1981) define 'understanding' in terms of pupils' ability to solve "problems ... recognisably connected to the mathematics curriculum but which ... require ... methods which [are] not obviously 'rules'" (Hart & Johnson, 1983, p.2). There are many nuances in these different approaches, but all highlight the importance of sense-making and of organising and connecting mathematical knowledge.

² Links between symbols and words for numbers (e.g., '5×5' and 'twenty-five') are largely associative and arbitrary. Number bonds and tables, if fluent, may be very densely conceptually embedded; e.g., rapid retrieval may involve some self-monitoring: for example, "9 7s are 56 – no that can't be right – 63".

³ It is important to emphasise that classroom talk is important to the development of conceptual knowledge and to doing mathematics in general. Hence, much classroom talk will be strategic and conceptual in nature.

⁴ We note that the relationship between metacognition and self-regulation is a current and disputed question, and researchers disagree on which is superordinate (see Gascoine et al., 2017).

⁵ By 'the gap', we mean the differences in understanding between the middle 95% of pupils in the age cohort (from the 2.5th to the 97.5th percentiles of attainment); i.e. two standard deviations either side of the mean.

4 Guide to Reading the Modules

4.1 Meta-analysis, effect sizes and systematic reviews

In this review, we have primarily drawn on *meta-analyses* rather than original studies. *Meta-analysis* is a statistical procedure for combining data from multiple studies. If a collection of studies are similar enough, and each reports an *effect size*, the techniques of meta-analysis can be used to find an *overall effect size* that indicates the best estimate of the underlying effect size for all of those studies.

In education, *effect size* (ES) is usually reported as Cohen's *d* or Hedges' *g*, which are measures of the difference between two groups in units determined by the standard deviation (the variation or spread) within the groups. An effect size of +1 means that the mean of the intervention group was 1 standard deviation higher than that of the control group. In practice, an effect size of 1 would be extremely large, and typical effect sizes of potential practical significance in education tend to be around the 0.1-0.5 range. Given our focus on experimental and quasi-experimental studies, we have largely reported measures of effect sizes using Cohen's *d* or Hedges' *g*. See Appendix: Technical (Section 14) for a definition of other measures of effect size reported or referred to in this review.

Caution should be exercised in comparing effect sizes for different interventions which may not be truly comparable in any meaningful way. Judgment is always required in interpreting effect sizes, and it may be more useful to focus on the order of related effect sizes (higher or lower than some other effect size) rather than the precise values. It should be noted that effect sizes are likely to be larger in small, exploratory studies carried out by researchers than when used under normal circumstances in schools. Effect sizes may be artificially inflated when the tests used in studies are specifically designed to closely match the intervention, and also when studies are carried out on a restricted range of the normal school population, such as low attainers, for whom the spread (standard deviation) will be smaller.

Where meta-analyses were not available in a particular area, we have instead made use of *systematic reviews*, which are a kind of literature review that brings together studies and critically analyses them, where computing an overall effect size is not possible, to produce a thorough summary of the literature relevant to a particular research question.

4.2 Structure

For each module, we give a *headline*, summarising the key points, followed by a description of the main *findings*. We summarise the *evidence base* from which this has arisen, and then comment on what we perceive to be the *directness* or relevance of the findings for schools in England. We score directness on a 1-3 scale of low-high directness on several criteria:

Where and when the studies were carried out: in some modules, the majority of the original studies were carried out in the United States, whilst in others many studies were conducted more than 25 years ago, and the directness score reflects our judgment of the extent to which the contexts, taken as a whole, are relevant to the current situation in England.

How the intervention was defined and operationalized: the extent to which the intervention or approach as described is the same as the intervention could be if adopted by teachers in England.

Any reasons for possible ES inflation: the extent to which the reported effect sizes may be artificially inflated.

Any focus on particular topic areas: the extent to which the findings about effectiveness of intervention or approach are relevant across mathematics as a whole.

Age of participants: in some modules, many of the original studies were conducted with older or younger learners, and the directness score reflects our judgment of extent to which the findings are relevant to the Key Stage 2 and 3 age group.

Finally, we provide details of the meta-analyses and other literature used.

5 Method

5.1 Our approach to analysing and synthesising the literature

Our approach was to carry out *second-order meta-analysis* – i.e., meta-analyses of existing meta-analyses – and occasionally third-order meta-analyses, where we summarise the findings of existing second-order meta-analyses. Second-order meta-analyses (also known as *umbrella reviews* or *meta-meta-analyses*) have been widely used in the medical and health sciences, and are becoming more frequent in educational research (Higgins, 2016). The intention of this set of second-order meta-analyses is to summarise the current evidence on teaching mathematics, as well as identify areas in which future meta-analyses and primary studies might be profitably directed.

We have not conducted a quantitative meta-analysis of any set of first-order meta-analyses. There were very few areas where several meta-analyses employed sufficiently similar research questions, theoretical frameworks and coding schemes to make a quantitative meta-analysis valid and straightforward to interpret. Instead, we present the results of the set of meta-analyses in tables, and we have adopted a narrative approach to synthesising the findings in each area. We have drawn on additional research when necessary to supplement the synthesis of the meta-analyses for each research question, particularly where the research evidence in a particular area is limited or the findings require interpretation or translation for the context in England. Where possible, we have drawn on recent high-quality systematic reviews, but, in some cases, where the evidence base is weak, we have taken account of research reporting single studies.

5.2 Limitations

Whilst our second-order meta-analytic approach has several advantages, there are disadvantages. We are dependent on the theoretical and methodological decisions that underpin the existing meta-analyses, and inevitably some nuance is lost in our focus on the “big picture”. We note also that there is an active debate on the statistical validity of meta-analytic techniques in education (Higgins & Katsipatakis, 2016; Simpson, 2016). Effect sizes are influenced by many factors, including research design, outcome measures or tests, and whether a teaching approach was implemented by the researchers who designed it or teachers. Meta-analyses of the highest quality use *moderator analysis* to examine whether these and other factors affect the magnitude of the effect sizes.

5.3 Data set

Our data set consists of 66 meta-analyses and 56 other relevant papers (mainly systematic reviews), written in English, relevant to the learning of mathematics of students aged 9-14, and published between 1970 and February 2017. These were identified using searches of electronic databases, the reference lists of the literature itself and our own and colleagues’ knowledge of the literature. See Sections 15 and 16 (Appendices: Literature Searches, and Inclusion / Exclusion Criteria) for further detail.

5.4 Coding and data extraction

Each paper was coded as a meta-analysis, systematic review or ‘other literature’, and details were recorded, including year of publication, author key words, abstract,

content area, main focus, secondary focus, key definitions, research questions, ranges of effect sizes, any pooled effect sizes and standard errors, number of studies and number of pupils, age range, countries studies conducted in, study inclusion dates, any pedagogic or methodological moderators or other analyses, inclusion/exclusion criteria and quality judgments. We assessed the methodological quality of the meta-analyses using six criteria, which we developed, informed by the PRISMA framework for rating the methodological quality of meta-analyses (<http://www.prisma-statement.org/>) and the AMSTAR criteria (Shea et al., 2009). For each meta-analysis, we graded each of our six criteria on a 1-3 (1 low, 3 high) scale.

The strength of evidence assessments were based on the GRADE system in medicine (Guyatt et al., 2008). This is an expert judgment-based approach that is informed, but not driven, by quantitative metrics (such as number of studies included). These judgements took account of the number of original studies, the methodological quality of the meta-analysis (including limitations in the approach or corpus of studies considered), consistency of results, the directness of results, any imprecision, and any reporting bias. Two members of the research team independently gave a high/medium/low rating for each section. Disagreements were resolved through discussion.

References

- Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/Pàngarau: Best evidence synthesis iteration*. Wellington, New Zealand: Ministry of Education.
- Conway, P. (2005). *International trends in post-primary mathematics education. Research report commissioned by the National Council for Curriculum and Assessment (NCCA)* Retrieved from <http://www.ncca.ie/uploadedfiles/mathsreview/intpaperoct.pdf>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924-926. doi: 10.1136/bmj.39489.470347.AD
- Higgins, S., & Katsipatakis, M. (2016). Communicating comparative findings from meta-analysis in educational research: some examples and suggestions. *International Journal of Research & Method in Education*, 39(3), 237-254. doi:10.1080/1743727X.2016.1166486
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding It Up: Helping Children Learn Mathematics (Prepared by the Mathematics Learning Study Committee, National Research Council)*. Washington DC: The National Academies Press.
- Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning*. London: Nuffield Foundation.
- Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., . . . Boers, M. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology*, 62(10), 1013-1020. doi:10.1016/j.jclinepi.2008.10.009
- Simpson, A. (2017). The misdirection of public policy: comparing and combining standardised effect sizes. *Journal of Education Policy*, 1-17. doi:10.1080/02680939.2017.1280183

6 Pedagogic Approaches

6.1 Feedback and formative assessment

What is the effect of giving feedback to learners in mathematics?

The general findings in the EEF toolkit on feedback appear to apply to mathematics: research tends to show that feedback has a large effect on learning, but the range of effects is wide and a proportion of studies show *negative* effects. The effect of formative assessment is more modest, but is more effective when teachers receive professional development or feedback is delivered through computer-assisted instruction. In mathematics, it may be particularly important to focus on the aspects of formative assessment that involve feedback. Feedback should be used sparingly and predominantly reserved for more complex tasks, where it may support learners' perseverance. The well-established literature on misconceptions and learners' understandings in mathematics provides a fruitful framework to guide assessment and feedback in mathematics.

Strength of evidence: HIGH

Definitions

In this review, feedback is conceptualised as “information provided by an agent (e.g. teacher, peer, book, parent, self, experience) regarding aspects of one’s performance or understanding” (Hattie & Timperley, 2007, p. 81). Formative assessment is broadly conceptualised as practices in which “information was gathered and used with the intent of assisting in the learning and teaching process” (Kingston & Nash, 2011, p. 29). Giving *feedback* means informing learners about their progress, whereas *formative assessment* refers to a broader process in which teachers clarify learning intentions, engineer activities that elicit evidence of learning and activate students as learning resources for one another as well give feedback (William & Thompson, 2007).

Findings

Feedback is generally found to have large effects on learning, and the Education Endowment Foundation Teaching and Learning Toolkit’s (EEF, 2017) second-order meta-analysis found an overall ES of $d=0.63$ on attainment across all subjects. There is considerable variability in reported effects, with some studies reporting negative effects, and Hattie and Timperley (2007) warn that feedback can have powerful negative as well as positive impacts on learning. Few of the existing meta-analyses on feedback examine mathematics specifically, but rather focus on the nature and causes of variability. However, many of the original studies are in the context of mathematics learning, and two meta-analyses report ESs for feedback in mathematics in comparison to other subjects: Scheerens et al. (2007) report that effects for mathematics ($d = 0.14$) are greater than for other subjects in general ($d = 0.06$), and similar to those for reading, whilst Bangert-Drowns et al. (1991) find no significant differences between subjects, although these differences may be related to the groups of subjects that are compared. Hence, the general findings in the toolkit on feedback would appear to apply to mathematics.

Kingston & Nash’s (2011) recent meta-analysis focuses on the wider strategy of formative assessment, of which feedback is a part, and their findings indicate a more

modest effect for formative assessment. Indeed, whereas feedback appeared to be particularly effective in mathematics, the opposite appears to be the case for formative assessment, with Kingston & Nash reporting an ES for mathematics of 0.17, compared with 0.19 for science and 0.32 for English Language Arts. This suggests that, in mathematics, it may be particularly important to focus on the aspects of formative assessment associated with feedback.

One meta-analysis suggests that feedback in mathematics is effective for low-attaining students ($d = 0.57$) (Baker et al., 2002), although this effect may be inflated due to the restricted attainment range of the population, and a further meta-analysis finds a lower, although still positive, effect for students with learning disabilities ($d = 0.21$) (Gersten et al., 2009).

It is important to understand how to give and use feedback in order for these effects to be realised. EEF (2017) note that giving feedback can be challenging. The evidence indicates that feedback should be clear, task-related and encourage effort (e.g., Hattie & Timperley, 2007). Feedback appears to be more effective when it is specific, highlights how and why something is correct or incorrect and compares the work to students' previous attempts (Higgins et al., 2017). Feedback is most likely to be beneficial if used sparingly and for challenging or conceptual tasks, where delayed feedback is beneficial (see Soderstrom & Bjork, 2013). The well-established literature on misconceptions and learner understandings in mathematics may provide a fruitful framework to guide assessment and feedback in mathematics (see Misconceptions module).

Kingston & Nash's (2011) analysis examined different ways in which formative assessment was implemented. Two approaches appeared to be more effective than others: one was based on professional development and the other was computer-based. These approaches yielded mean effect sizes of 0.30 and 0.28 respectively. In comparison, other approaches, such as curriculum-embedded formative assessment systems, which "involved administering open-ended formative assessments at critical points throughout the curriculum in order to gain an understanding of the students' learning processes" (p. 32), had nil or very small effects.

Evidence base

We have drawn on four meta-analyses providing recent evidence of the impact of feedback in mathematics specifically. These synthesise a total of 275 studies with the date range 1982-2010. The four meta-analyses are all judged to be of medium or high methodological quality. While overall there is noted to be wide variability in studies looking at the effect of feedback across subjects, the ESs reported in these meta-analyses for mathematics are fairly consistent, with the exception of Baker et al. (2002), where the higher ESs may be accounted for by the inclusion of studies involving computer feedback.

There is a need for more research on the nature of feedback specifically in mathematics. Kingston & Nash (2011) argue that, with formative assessment practices (which include feedback) in wide use, and with the potential of them to produce high effects, the paucity of the current research base is problematic.

Meta-analysis	Focus	<i>k</i>	Quality	Date Range
---------------	-------	----------	---------	------------

Baker et al. (2002)	Instructional strategies in mathematics for low-achieving students	15	2	1982-1999
Gersten et al. (2009)	Instructional strategies in mathematics for students with learning difficulties	41	3	1982-2006
Kingston & Nash (2011)	Formative assessment	42	3	1990-2010
Scheerens et al. (2007)	Review of the effectiveness of school-level and teaching-level initiatives	177	2	1995-2005

Directness

Overall we would assess the evidence base as being of high directness to the English context.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Studies were conducted in many countries, although a significant proportion were located in the US / UK. Scheerens et al. (2007) conducted a moderator analysis using 'country' of study as a variable and found results across countries to be broadly similar.
How the intervention was defined and operationalised	3	Feedback is generally clearly operationalised, although, as noted by Hattie & Timperley (2007) and others, feedback is not a straightforward strategy to implement and can have powerful negative as well as positive effects.
Any reasons for possible ES inflation	3	Two of the four meta-analyses looked at low-achieving learners or those with a learning disability and, in these cases, the effects may be inflated due to restricted samples.
Any focus on particular topic areas	3	
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Impact of providing feedback to students on mathematics attainment for students with learning disabilities</i>			

Gersten et al. (2009)	0.21 [0.01, 0.40]	12	<p>This study looked at interventions for LD students only.</p> <p>This effect size was calculated through the combination of student feedback ($g=0.23$ [0.05, 0.40], $k=7$) and goal-setting student feedback ($g=0.17$ [-0.15, 0.49], $k=5$).</p>
<i>Impact of providing feedback to teachers on mathematics attainment for students with learning disabilities</i>			
Gersten et al. (2009)	0.23 [0.05, 0.41]	10	
<i>Impact of providing feedback on mathematics attainment for low attaining students</i>			
Baker et al. (2002)	0.57 [0.27, 0.87]	5	<p>This study looked at interventions for low-achieving students only. In some cases this feedback was computer-generated (these studies are not segregated).</p> <p>The comparison group in these four studies either was provided with no performance feedback or with such limited feedback that a relevant contrast between the experimental and comparison group was meaningful.</p> <p>This is a moderate effect and the second largest mean effect size found in this synthesis.</p>
<i>Impact of providing feedback on mathematics attainment in comparison to other subjects</i>			
Scheerens et al. (2007)	0.136	152 included in moderator analysis across all subjects,	Coefficient from moderator analysis regression reported. Feedback is a broad category that includes monitoring, assessment, and tests.

		number of maths studies not reported; we would estimate this to be between 5 and 20.	Language ES = 0.143 All subjects ES = 0.06
<i>Impact of formative assessment on mathematics attainment in general</i>			
Kingston & Nash (2011)	0.17 [0.14, 0.20]	19	Moderator analysis showed content area had the greatest impact on mean effects. English language arts d= 0.32 [0.30, 0.34.] Science d= 0.09 [-0.09, 0.25]
<i>Impact of formative assessment focused professional development programmes on overall attainment</i>			
Kingston & Nash (2011)	0.30 [0.20, 0.40]	23	Studies were coded as PD where they examined “professional development that involved educators spending a period of time learning and focusing on how to implement various aspects of formative assessment techniques (e.g., commen- only marking, self-assessment, etc.) in their classrooms” (pp. 31-2)
<i>Impact of the use of a computer-based formative assessment system on overall attainment</i>			
Kingston & Nash (2011)	0.28 [0.26, 0.30]	6	Studies coded in this category “involved the online administration of short indicator level tests that provided score reports to teachers and are similar to state-wide assessments ... One of these systems incorporated an additional tutoring feature in the form

			of student-level scaffolding.” (p. 32)
--	--	--	--

References

Meta-analyses included

Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, 103(1), 51-73. doi:10.1086/499715<http://www.jstor.org/stable/1002308>

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

Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28-37. doi:10.1111/j.1745-3992.2011.00220.x

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). Review and meta-analyses of school and teaching effectiveness. Enschede: Department of Educational Organisation and Management, University of Twente.

Secondary meta-analyses included

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61(2), 213-238. doi:doi:10.3102/00346543061002213

Higgins, S., Katsipatakis, M., Villanueva-Aguilera, A. B. V., Coleman, R., Henderson, P., Major, L. E., Coe, R., & Mason, D. (2016). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Feedback*. London: Education Endowment Foundation.

<https://educationendowmentfoundation.org.uk/resources/teachinglearning-toolkit>

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.

Kluger, A. N., & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), 254-284.

Meta-analyses excluded

Browder, D. M., Spooner, F., Ahlgrim-DeLzell, L., Harris, A. A., & Wakemanxya, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional Children*, 74(4), 407-432. [Feedback here involves prompting and support fading – it is not feedback as conceptualised in mainstream classrooms]

Other references

Soderstrom, N. C., & Bjork, R. A. (2013). Learning versus performance. In D. S. Dunn (Ed.), *Oxford bibliographies online: Psychology*. New York, NY: Oxford University Press.

William, D., & Thompson, M. (2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

6.2 Collaborative learning

What is the evidence regarding the effect of using collaborative learning approaches in the teaching and learning of maths?

Collaborative Learning (CL) has a positive effect on attainment and attitude for all students, although the effects are larger at secondary. The largest and most consistent gains have been shown by replicable structured programmes lasting 12 weeks or more. Unfortunately, these programmes are designed for the US educational system, and translating the programmes (and the effects) for the English educational system is not straightforward. The evidence suggests that students need to be taught *how* to collaborate, and that this may take time and involve changes to the classroom culture. Some English-based guidance is available.

Strength of evidence: HIGH

Findings

The meta-analyses present definitions of CL ranging from the non-specific working with or among peers within group settings (Lee, 2000) through definitions built on Slavin's studies (e.g. Slavin, 2007, 2008), where "students of all levels of performance work together in small groups toward a common goal" (Othman, 1996, p. 10). Haas (2005) includes a far broader range of approaches, including whole-class collaboration, although this definition sits outside of the other literature. CL may co-occur with other approaches (such as peer-tutoring), with Reynolds & Muijs (1999, p. 238) suggesting that CL should be used alongside whole-class interactive approaches to produce "an optimal level of achievement across a range of mathematical skills". Furthermore, CL is commonly associated (particularly in the US) with specific programmes and approaches, such as Student Teams Achievement Divisions (STAD), Team Assisted Individualization (TAI), and dyadic methods (such as peer-tutoring). The five meta-analyses central to this evidence focussed on one or more of these programmes/approaches, although the majority of studies synthesised involved one particular programme, STAD.

The impact of CL on mathematics attainment reported within four meta-analyses ranged from an ES of 0.135 (Stoner, 2004) to 0.42 (Slavin et al., 2008). Slavin et al.'s finding, which is based on Middle and High School students, is higher than their finding (Slavin & Lake, 2007) for Elementary age students (0.29). In both cases, Slavin & Lake and Slavin et al. found CL, categorised together with other "innovative teaching approaches", to be among the most effective programmes. The finding of a higher effect size with older students aligns with Othman's (1996) moderator analysis, which also found a higher ES for secondary grades (0.29 compared with 0.18 for elementary). Slavin and his colleagues focused on replicable intervention programmes lasting 12 weeks or more, and found a larger effect than Othman for both Elementary, and Middle and High School, students. This suggests that students need to learn how to collaborate effectively.

Meta-analyses we judged as secondary to this overall analysis (Chen, 2004; Lee, 2000), on the basis of addressing a specific student population (lower-attainers and those with learning difficulties), suggest that CL may be less effective for this specific population. The needs of this population are considered in the module on responding to different attainment levels.

The impact of CL on attitudes to mathematics was reported within two meta-analyses, which found ESs of 0.20 (Othman, 1996) and 0.35 (Savelsbergh et al., 2016). Savelsbergh et al. found that, unlike attainment, the effect of CL on attitudes *decreases* as students got older (although we note that this may partly reflect the general trend that student attitudes to mathematics decrease with age).

Evidence base

Having excluded some meta-analyses due to their poor methodological quality and noted others as secondary to our evidence because of the population included, we found four meta-analyses examining the impact of CL on mathematics attainment for the general population, synthesizing a total of 79 studies over the period 1970–2003. Due to Slavin’s focus on robust studies of replicable intervention programmes lasting 12 weeks or more, the degree of study overlap was minimal [just three (14%) of Stoner’s (2004) and two (5%) of Othman’s (1996) included studies overlapping with the studies included across both of Slavin’s analyses], leading to our judgement that the strength of the evidence base is high.

We found two meta-analyses examining the impact of CL on mathematics attitude for the general population, which synthesised a total of 29 studies over the period 1970–2014. There is no overlap between the studies included in these meta-analyses.

All five included meta-analyses were rated as medium or high methodological quality. The range of reported effects is small. While the evidence for attitudes is more limited, these studies are fairly consistent in their findings.

Given the US focus of the majority of studies and programmes, there is a need for experimental research to evaluate the effects of interventions adapted or designed for English mathematics classrooms.

Meta-analysis	Focus	<i>k</i>	Quality	Date Range
Haas (2005)	The effect of CL on the learning of algebra	3	2	1980-2002
Othman (1996)	The effect of CL on mathematics attainment and attitude across Grades K-12	39	2	1970-1990
Savelsbergh et al. (2016)	The impact of different teaching approaches – including CL – on student attitudes in mathematics and science across Grades 3-11	5	3	1988-2014
Slavin and Lake (2007)	The impact of a range of replicable programmes lasting 12 weeks or more – including CL programmes – on Elementary mathematics achievement	9	3	1985-2002
Slavin et al. (2008)	The impact of a range of replicable programmes lasting 12 weeks or more – including CL programmes –	9	3	1984-2003

	on Middle and High School mathematics achievement			
Stoner (2004)	The effect of CL on mathematics attainment in the middle grades	22	2	1972-2003

Directness

Our overall judgement is that the available evidence is of medium directness.

The majority of the programmes examined in these meta-analyses are set in the US and, inevitably, the programmes are designed around the particularities of the US school system. Translating an intervention programme from one system to another is not straightforward, particularly where, as with CL, a programme is designed to alter the social norms of the mathematics classroom. The recent UK trial of PowerTeaching Maths (Slavin et al, 2013) demonstrates this difficulty. PowerTeaching Maths is a technology-enhanced teaching approach based around co-operative learning in small groups. However, the effects found in US experimental studies were not replicated in the UK. The researchers found that implementation was limited by the prevalence of within-class ability grouping in England, which appeared to affect teachers' implementation of key aspects of the approach.

Nevertheless, the evidence from US programmes does suggest that the success of CL interventions relies on a structured approach to collaboration and that students need to be taught how to collaborate. Some UK-focused interventions have shown positive effects in quasi-experimental studies, such as the SPRinG approach in KS1, 2 and 3, for which teacher guidance is readily available, although this is not specific to mathematics (Baines et al., 2014). Evidence-based guidance on CL at secondary is readily available in schools (Swan, 2015).

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	All meta-analyses were US-based and few of the included studies were located in the UK. These meta-analyses predominantly considered specific CL programmes rather than a more general notion of CL which may be applied in the UK. Slavin et al. (2013) note that extensive professional development is a common feature to these programmes given to teachers embarking on such approaches, suggesting that the positive impacts of US CL approaches "can be readily disseminated."
How the intervention was defined and operationalised	3	CL clearly defined and usually associated with specific programmes.
Any reasons for possible ES inflation	3	No – meta-analyses related to LA and LD populations taken out of main analysis.
Any focus on particular topic areas	3	
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Effect of Collaborative Learning (CL) on mathematical attainment</i>			
Othman (1996): all grades	0.266	39	Range of CL approaches included: Students Teams-Achievement Division [STAD], Team Assisted Individualization [TAI], Teams-Games-Tournament [TGT], Learning Together, Peer Tutoring
Stoner (2004); middle grades	0.135	22	Included range of general CL approaches and specialist programmes including: STAD, TAI and TGT
<i>Effect of replicable Collaborative Learning (CL) programmes lasting 12 weeks or more on mathematical attainment</i>			
Slavin and Lake (2007): primary	0.29	9	Examined three US CL programmes: Classwide Peer Tutoring, Peer-Assisted Learning (PALS), STAD
Slavin et al. (2008): secondary	0.42	9	Examined four US CL programmes: STAD, PALS, Curriculum-Based Measurement, IMPROVE
<i>Effect of Collaborative Learning (CL) on attitudes to mathematics</i>			
Othman (1996): all grades	0.20	24	
Savelsbergh et al. (2016): all grades	0.35	5	95% CI [0.24; 0.47] Total of 65 experiments from 56 studies. Only 5 looked at CL in maths
<i>Effect of Collaborative Learning (CL) on learning of algebra</i>			
Haas (2005)	0.34	3	

References

Meta-analyses included

- Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin*, 89(642), 24-46.).
- Othman, N. (1996). The effects of cooperative learning and traditional mathematics instruction in grade K-12: A meta-analysis of findings. Doctoral Thesis, West Virginia University. ProQuest UMI 9716375
- Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, 19, 158-172.
- Slavin, R. E. and Lake, C. (2007). *Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis* (Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University, 2007).
- Slavin, R. E., Lake, C., & Groff, C. (2008). *Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis* (Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University, 2007).
- Stoner, D. A. (2004). The effects of cooperative learning strategies on mathematics achievement among middlegrades students: a meta-analysis. University of Georgia, Athens.

Secondary Meta-analyses included

- Chen, H. (2004). The efficacy of mathematics interventions for students with learning disabilities: A meta-analysis. (Order No. 3157959, The University of Iowa). ProQuest Dissertations and Theses, 136 pages. Retrieved from <http://search.proquest.com/docview/305196240?accountid=14533>. (305196240).
- Lee, D. S. (2000). A meta-analysis of mathematics interventions reported for 1971-1998 on the mathematics achievement of students identified with learning disabilities and students identified as low achieving. Doctoral Thesis, University of Oregon ProQuest UMI 9963449

Meta-analyses excluded

- Capar, G., & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory & Practice*, 2, 553-559. [Weak methodology as evidenced by apparent errors and other anomalies in the data extraction and search criteria]
- Nunnery, J.A., Chappell, S. & Arnold, P. (2013). A meta-analysis of a cooperative learning model's effects on student achievement in mathematics. *Cypriot Journal of Educational Sciences*, 8(1), 34-48. [Some reporting issues and a high degree of overlap with Slavin's studies]
- Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). *Methods of Instructional Improvement in Algebra A Systematic Review and Meta-*

Analysis. *Review of Educational Research*, 80(3), 372-400. [CL is included as a moderator variable but analysis not extended]

Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review*, 22(4), 393-409. [Examines worked examples presented in various ways rather than cooperative learning and its outcomes per se]

Other references

Baines, E., Blatchford, P., Kutnick, P., Chowne, A., Ota, C., & Berdondini, L. (2014). *Promoting Effective Group Work in the Primary Classroom: A handbook for teachers and practitioners*. London: Routledge.

Reynolds, D. and Muijs, D. (1999). The effective teaching of mathematics: a review of research. *School Leadership & Management*, 19(3), 273-288.

Slavin, R. E. (1999). Comprehensive approaches to cooperative learning. *Theory into Practice*, 38(2), 74-79.

Slavin, R. E. (1984). Team-Assisted individualization; Cooperative learning and individualized instruction in the mainstreamed classroom. *Remedial and Special Education*, 5, 33-42.

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, 50, 315-342.

Slavin, R. E., Sheard, M., Hanley, P., Elliott, L., & Chambers, B. (2013). *Effects of Co-operative Learning and Embedded Multimedia on Mathematics Learning in Key Stage 2: Final Report*. York: Institute for Effective Education.

Swan, M. (2005). *Improving learning in mathematics*. London: Department for Education and Skills.

6.3 Discussion

What is known about the effective use of discussion in teaching and learning mathematics?

Discussion is a key element of mathematics teaching and learning. However, there is limited evidence concerning the effectiveness of different approaches to improving the quality of discussion in mathematics classrooms. The available evidence suggests that teachers need to structure and orchestrate discussion, scaffold learners' contributions, and develop their own listening skills. Wait time, used appropriately, is an effective way of increasing the quality of learners' talk. Teachers need to emphasise learners' explanations in discussion and support the development of their learners' listening skills.

Strength of evidence: LOW

Introduction

Discussion is an important tool for learning mathematics. However, there is limited evidence concerning the effectiveness of different approaches aimed at improving the quality of discussion in mathematics classrooms. We found no meta-analyses looking at discussion in mathematics, and only three systematic reviews.

Findings

Effective discussion in the mathematics classroom goes beyond setting up opportunities for talk. Eliciting and supporting effective dialogue is not simple (Walshaw & Anthony, 2008). Much classroom discourse follows the initiation-response-evaluation (IRE) model, in which the teacher *initiates* by asking a question, the learner *responds* by answering the question and the teacher then gives an *evaluation*. While this has its uses, classroom discussion can be enhanced by facilitating more extended contributions from all learners (Kyriacou & Issitt, 2008). Alexander et al. (2010) argue that *dialogic teaching* is crucial to advancing learning. In contrast to IRE, dialogic teaching involves a back-and-forth between the learners and the teacher, and requires careful and effective structuring (Alexander, 2017). The classroom culture and the actions of the teacher need to allow all learners to contribute equally; Walshaw and Anthony (2008) cite a number of studies suggesting that particular students often dominate discussion in the mathematics classroom.

Increasing *wait time*, the time a teacher pauses after asking a question before accepting learner responses, has been shown to be an effective way of increasing the quality of talk (Tobin, 1986, 1987). Wait time in mathematics lessons is typically less than 1 second, suggesting that priority is often given to maintaining a brisk pace with a focus on quickly obtaining correct 'answers'. Evidence suggests that increasing wait time to around 3 seconds, particularly when higher-order questions are used, can have dramatic effects on learners' involvement in classroom discussion, leading to higher-quality responses from a greater range of learners. A further increase of wait time to more than 5 seconds, however, *decreases* the quality of classroom talk (Tobin, 1987).

Improving mathematics dialogue is more complicated than just instigating 'more talk'; *effective* talk also requires effective listening, particularly so on the part of the teacher (Kyriacou & Issitt, 2008). Teachers need to listen actively to learners' contributions, particularly their explanations, and show genuine interest in these,

rather than listening in an evaluative manner for expected answers (Walshaw & Anthony, 2008). The focus of talk needs to shift from evaluation (judging the correctness of an answer) to exploration of mathematical thinking and ideas (Kyriacou & Issitt, 2008; Walshaw & Anthony, 2008). Teachers need to teach learners how to discuss and “what to do as a listener” (Walshaw & Anthony, 2008, p. 523). Walshaw and Anthony (2008) cite studies finding that some primary learners simply do not know how to explain mathematical ideas, and that the teacher needs to establish norms for what counts as mathematically acceptable explanation.

Effective discussion is likely to be part of collaborative approaches to learning. This will include elements of listening, reflection, evaluation, and self-regulation (Kyriacou & Issitt, 2008). Discussing mathematics can help to make learners’ thinking visible and enable ideas to be critiqued (Walshaw & Anthony, 2008).

Evidence base

As stated above, the evidence base examining discussion in mathematics is limited. We identified no relevant meta-analyses and, hence, we draw on three research syntheses.

Research-synthesis	Focus and core findings
Kyriacou & Issitt (2008)	<p>UK study of mathematics lessons Covered Key stages 2 and 3 (learners aged 7 – 14) Analysis of 15 primary studies Examined the characteristics of effective teacher-initiated teacher-pupil dialogue Focussed on outcome measure of conceptual understanding in mathematics Noted the dominance of IRE and the need to go beyond this Strongest evidence came from studies in which teachers taught learners how to make use of dialogue Identified paucity of evidence in the area</p>
Tobin (1987) See also Tobin (1986)	<p>Australian review of studies involving wait time in a range of subject areas and grade levels Identified 6 primary studies in which wait-time was not manipulated: <ul style="list-style-type: none"> ○ 4 studies included learners aged 9-14 ○ 2 studies involved mathematics Identified 19 primary studies in which wait-time was manipulated: <ul style="list-style-type: none"> ○ 13 studies included learners aged 9-14 ○ Only 1 study involved mathematics (68% were in science) Found that a wait time of longer than 3 seconds resulted in changes to teacher and student discourse</p>

	<p>Suggests that the additional ‘think time’ may result in higher cognitive learning</p> <p>Cautions against the simplistic notion of increasing wait-time to make classrooms more effective</p>
Walshaw & Anthony (2008)	<p>New Zealand review of primary studies into how teachers manage discourse in mathematics classrooms</p> <p>Draws on the data set of Anthony & Walshaw’s (2007) <i>Effective Pedagogy in Mathematics/Pangarau: Best Evidence Synthesis Iteration</i> (see references elsewhere in this review)</p> <p>Theorises mathematics classrooms as activity systems in understanding effective pedagogy (in relation to dialogue)</p> <p>Four core requirements:</p> <ol style="list-style-type: none"> A classroom culture where all learners are able to participate equally Ideas are coproduced through dialogue, extending other learners’ thinking Teachers do not accept all answers but listen attentively and help to build dialogue to develop mathematical ideas Teachers need the subject knowledge and flexibility to spot, help learners make sense of, and develop, mathematically grounded understanding

Directness

While the available evidence is limited, that which we found has direct relevance to the English mathematics classroom context.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	The three reviews were conducted in the UK or Australasia, drawing on a range of primary studies. The conclusions have applicability to the English mathematics classroom context.
How the intervention was defined and operationalised	3	All three reviews carefully define dialogue / wait-time.
Any focus on particular topic areas	2	Kyriacou & Issitt (2008) and Walshaw & Anthony (2008) focus solely on mathematics. Some caution should be exercised in applying the findings from Tobin’s (1987) review of wait time to the mathematics classroom.

Age of participants	3	Large crossover with our focus on learners aged 9-14.
---------------------	---	---

References

Research syntheses

- Kyriacou, C. and Issitt, J. (2008) What characterises effective teacher-initiated teacher-pupil dialogue to promote conceptual understanding in mathematics lessons in England in Key Stages 2 and 3: a systematic review. Technical report. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Tobin, K. (1987). The Role of Wait Time in Higher Cognitive Level Learning. *Review of Educational Research*, 57(1), 69-95. doi:10.3102/00346543057001069
- Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, 78(3), 516-551.

Other references

- Alexander, R. J. (2017). *Towards dialogic teaching: Rethinking classroom talk*. (5th edition) York: Dialogos.
- Alexander, R.J. (ed) (2010) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review*. London: Routledge.
- Tobin, K. (1986). Effects of teacher wait time on discourse characteristics in mathematics and language arts classes. *American Educational Research Journal*, 23(2), 191-200.

6.4 Explicit teaching and direct instruction

What is the evidence regarding explicit teaching as a way of improving pupils' learning of mathematics?¹

Explicit instruction encompasses a wide array of teacher-led strategies, including direct instruction (DI). There is evidence that structured teacher-led approaches can raise mathematics attainment by a sizeable amount. DI may be particularly beneficial for students with learning difficulties in mathematics. But the picture is complicated, and not all of these interventions are effective. Furthermore, these structured DI programmes are designed for the US and may not translate easily to the English context. Whatever the benefits of explicit instruction, it is unlikely that explicit instruction is effective for all students across all mathematics topics at all times. How the teacher uses explicit instruction is critical, and although careful use is likely to be beneficial, research does not tell us how to balance explicit instruction with other more implicit teaching strategies and independent work by students.

Strength of evidence: MEDIUM

Findings

Explicit instruction refers to a wide array of “teacher-led” approaches, all focused on teacher demonstration followed by guided practice and leading to independent practice (Rosenshine, 2008). Explicit instruction is not merely “lecturing”, “teaching by telling” or “transmission teaching”. Although explicit instruction usually begins with detailed teacher explanations, followed by extensive practice of routine exercises, it later moves on to problem-solving tasks. However, this always takes place *after* the necessary ideas and techniques have been introduced, fully explained and practised, and not before. In this way, explicit instruction differs from inquiry-based learning or problem-based learning approaches, in which, typically, students are presented (for example, at the *start* of a topic) with a problem that they are not expected to have any methods at their fingertips to solve (Rosenshine, 2012).

A very important and the most heavily-researched example of explicit instruction is *direct instruction* (DI), which exists in various forms. *Direct Instruction* (with initial capital letters, here always written in italics) refers to a particular pedagogical programme, first developed by Siegfried Engelmann in the US in the 1960s. This was designed to be implemented as a complete curriculum, and involves pre- and post-assessments to check students' readiness and mastery, teacher scripts, clear hierarchies of progression, a fast pace, breaking tasks into small steps, following one set approach and positive reinforcement. Looser understandings of DI than this draw on some of these features without adopting the full programme in its entirety. At its core, DI stresses the modelling of fixed methods, explaining how and when they are used, followed by extensive structured practice aimed at mastery. (Note that this understanding of ‘mastery’ is different from mastery as currently being promoted in England, although it has some similarities to Bloom's [1968] approach to mastery – see the Mastery module, 6.5.)

There is strong evidence for medium to high effects of both DI in general and *Direct Instruction* in particular on mathematics attainment (e.g., Dennis et al., 2016;

¹ For an English audience, we have chosen to refer to ‘explicit teaching’ in the title question, although the research literature refers in the main to ‘explicit instruction’.

Gersten et al., 2009), and some evidence that DI is particularly beneficial for students with learning difficulties in mathematics (e.g., Chen, 2004; Haas, 2005). However, a large range of effect sizes has been reported (for example, from 0.08 to 2.15 in Gersten et al., 2009). It is possible that some of the “too good to be true” effect sizes have been inflated due to methodological features of the studies (see below).

There is some indication that when teacher instruction is more explicitly given, and students’ activity is more tightly specified, larger effects on attainment are obtained (Gersten et al., 2009). Gersten et al. (2009) contrasted L. S. Fuchs, Fuchs, Hamlett, et al. (2002), who reported an effect size of 1.78 for students who were taught to solve different word problems step by step, with Ross and Braden (1991), where the effect size was 0.08, and where students worked through “reasonable steps to solve the problem but are not explicitly shown how to do the calculations” (p. 1216). It may be that the more tightly focused the DI is on the procedure or concept being learned, the higher the ES.

As discussed in Section 3, almost every strategy benefits from the judicious use of “explicit instructional guidance” in some form. Jacobse & Harskamp (2011, p. 26) found that DI had an effect of similar size to other “constructivist” strategies, including “guided discovery”. In contrast to this, there is considerable evidence that “pure” *unguided* discovery (unstructured exploration) is less effective [Mayer, 2004; see also Askew et al.’s (1997) study of effective teaching of numeracy in primary schools in England, which found that effective teachers tended to have a connectionist rather than a transmissionist or a discovery orientation towards teaching mathematics]. The literature on DI does not address the question of how to balance explicit teaching with other less “direct” teaching strategies and independent work by students.

Explicit instruction has been criticised by some as an excessively regimented approach (Borko & Wildman, 1986) with an undesirable focus on rote factual knowledge and preparation for tests, with students in a passive learning mode (Brown & Campione, 1990) and teachers reduced, in some cases, to merely reading out a script. However, proponents of forms of explicit instruction argue that creating an instructional sequence that is carefully based on research allows students’ skills to be sequenced, so that they learn in a cumulative and efficient way (McMullen & Madelaine, 2014).

Horak (1981) found no overall effect for individualised instruction in comparison to traditional instruction.

Evidence base

We identified seven meta-analyses synthesising a total of 126 unique studies. Three of these meta-analyses were of overall high quality, although we had reservations about aspects of the methodologies – in particular, lack of clarity over definitions of explicit instruction and possible biases associated with search and inclusion criteria. The other four meta-analyses were of medium quality. Pooled effect sizes across the meta-analyses ranged from 0.55 to 1.22.

As mentioned above, Gersten et al. (2009) found a large range of effect sizes for DI, from 0.08 to an enormous 2.15. It is possible that some of the high effect sizes could have been obtained as a consequence of interventions being used specifically with low-attaining subsets of the population (which have smaller standard deviations,

leading to inflated effect sizes) or as a result of regression to the mean when selecting study participants based on previous low attainment. It may also be the case that the “directness” of DI approaches makes these inherently more likely to produce high effect sizes, since the match between the intervention and the post-test is likely to be high for an intervention which explicitly tells students what they are supposed to be learning. It is arguable to what extent this constitutes fair measurement of the intervention or is an artefact of the style of this particular kind of intervention (Haas, 2005). Very few studies included delayed post-tests, which would help to assess longer-lasting effects of explicit instruction. The specific focus of tests used is also important; we would expect higher effect sizes where tests related to precisely the method being taught, but if learners were tested on their ability to transfer their knowledge to some related but different problem, it could be that explicit instruction approaches would be found to be less effective.

Gersten et al.’s (2009) pooled effect size of 1.22 might be regarded as inflated, since it is well outside the normal range of effect sizes obtained for educational interventions. As mentioned above, the range in Gersten et al. (2009) is very large (0.08 to 2.15), with a Q statistic of 41.68 (df = 10, $p < .001$), meaning that it is not reasonable to suppose that there is a single true underlying effect size for these studies.

Gersten et al. (2009) have reservations regarding the methodology used by Kroesbergen & Van Luit (2003) in finding that DI and self-instruction were more effective than mediated instruction. Reported ESs from small (or even single-subject) designs (or those focused exclusively on very low attainers) may not be reliable indicators of likely gains in terms of the entire cohort.

The percentages of overlapping studies between the meta-analyses used here are generally small, except for Baker et al. (2002), Gersten et al. (2009) and Kroesbergen & Van Luit (2003). All of the other meta-analyses have percentages of unique studies (not shared with any of the other meta-analyses) over 60%. This could be a result of different definitions of DI leading to different subsets of studies being selected.

Directness

DI has been strongly promoted as a highly effective approach to teaching (e.g., Gersten, Baker, Pugach, Scanlon, & Chard, 2001). The majority of the studies synthesised were carried out in the US. General similarities between the school systems in England and the US contribute to the directness of these findings. However, the ‘social validity’ of an intervention such as DI could be weak in England, where teachers tend to be less comfortable with teacher-centred and highly directed approaches than they may be in the US. (See the “Textbooks” module – 7.5 – for further detail.)

A very large proportion of the DI studies synthesised in meta-analyses are with low-attaining students. Not only is this potentially problematic in terms of inflated effect sizes (as discussed above), but it also threatens the directness of these findings for generalisability to the whole cohort.

The variety of definitions of DI is also highly problematic, as it is sometimes unclear that like is being compared with like, both within a single meta-analysis but, even more so, when bringing together several different meta-analyses carried out by different authors. Some studies combine DI interventions which appear to vary

considerably. In Baker et al. (2002), for instance, two of the four studies are Engelmann-influenced *Direct Instruction* studies, both with very small samples ($N = 35$ and 29), which feature video instruction as well as teacher-led instruction. However, the overall ES reported by Baker et al. is driven by the other two studies, which are based on the Mayer (2004) heuristic method for problem-solving ($N = 90$ and 489).

Implementation of DI in mathematics in England would have to take account of numerous factors, including content area, curriculum and resources. It would also be important to know for what length of time DI would need to be implemented for effects to be seen. In Haas (2005, p. 30), the mean length for interventions was about 11 weeks, and it could be that extended use of DI is necessary for sizeable effects to be seen. Whatever the benefits of DI, it is likely that DI is not equally effective for all students at all times.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	The majority of the studies were carried out in the US, but the conclusions have applicability to the English mathematics classroom context. However, there could be a 'social validity' problem with DI.
How the intervention was defined and operationalised	1	Varied definitions of DI.
Any focus on particular topic areas	1	Mainly low-attaining students.
Age of participants	3	

Further research

There is a need for research into DI approaches in England that includes delayed post-tests as well as investigation of the relative benefits for different topics and procedural versus conceptual learning. It is also important to explore the effect of different kinds of tests – those focused on far transfer from the context of the teaching would be particularly valuable. It would also be beneficial to have smaller-scale experimental studies before large-scale trials. An extensive theoretically-informed meta-analysis and a systematic review are both needed.

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Quality	Comments
Baker, Gersten & Lee (2002)	0.58	4	2	Broad review on teaching mathematics to low-achieving students included studies coded as "explicit instruction": "In these studies, the manner in which concepts and problem solving were taught to students was far more explicit than is typical." (p. 63)

Chen (2004)	1.01	8	3	<p>Focused on mathematics interventions for students with learning disabilities. They see DI as “based on teacher-led, structured, and systematic explicit instruction” (p. 4) “The characteristics of direct instruction highlight fast-paced, well-sequenced, highly focused lessons, delivering lessons in a small-group, providing ample opportunities for students to respond and instant corrective feedback” (p. 19). They state confidently that “it is safe to conclude that direct instruction is highly effective for mathematics remediation for students with learning disabilities.” (p. 108)</p>
Dennis et al. (2016)	0.76	18	3	<p>Use Baker et al.’s classification and found explicit teacher-led instruction to be the second most effective approach that they looked at (following peer-assisted learning).</p>
Gersten et al. (2009)	1.22	11	3	<p>They included studies if all three of these criteria were met:</p> <p>(a) The teacher demonstrated a step-by-step plan (strategy) for solving the problem, (b) this step-by-step plan needed to be specific for a set of problems (as opposed to a general problem-solving heuristic strategy), and (c) students were asked to use the same procedure/steps demonstrated by the teacher to solve the problem.</p> <p>Studies covered “a vast array of topics” (p. 1216).</p>
Haas (2005)	0.55	10	2	<p>Looked at secondary algebra. They define DI as “Establishing a direction and rationale for learning by relating new concepts to previous learning, leading students through a specified sequence of instructions based on predetermined steps that introduce and reinforce a concept, and providing students with practice and feedback relative to how well they are doing.”</p>

				Found that DI had the largest effect for low-ability and high-ability students (p. 30).
Horak (1981)	-0.07	129	2	Found a great deal of variation across individualised instruction approaches.
Jacobse & Harskamp (2011)	0.58	40	2	<p>Looked at effects of instructional interventions on students' mathematics achievement.</p> <p>Although the ES is 0.58 their finding was that there is no difference between direct and "indirect" instruction (ES = 0.61). However, they equate "indirect" with "the constructivist approach of guiding students instead of leading them" (p. 26). Their definition was "Direct instruction is an instructional approach where a teacher explicitly teaches students learning strategies by modeling and explaining why, when, and how to use them." (p. 5). They speculate that for students of low ability, DI may be most effective (p. 24), but cannot confirm this (p. 26).</p>
Kroesbergen & Van Luit (2003)	0.91	35	2	<p>Looked at elementary students with special needs (students at risk, students with learning disabilities, and low-achieving students) and examined a range of interventions.</p> <p>DI and self-instruction were found to be more effective than mediated instruction.</p>

References

Meta-analyses included

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

Chen, H. (2004). The efficacy of mathematics interventions for students with learning disabilities: A meta-analysis. (Order No. 3157959, The University of Iowa). ProQuest Dissertations and Theses, 136 pages. Retrieved from <http://search.proquest.com/docview/305196240?accountid=14533>. (305196240).

Dennis, M. S., Sharp, E., Chovanec, J., Thomas, A., Burns, R. M., Custer, B., & Park, J. (2016). A Meta-Analysis of Empirical Research on Teaching Students

with Mathematics Learning Difficulties. *Learning Disabilities Research & Practice*, 31(3), 156-168.

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

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin*, 89(642), 24-46.

Horak, V. M. (1981). A meta-analysis of research findings on individualized instruction in mathematics. *The Journal of Educational Research*, 74(4), 249-253. <http://dx.doi.org/10.1080/00220671.1981.10885318>

Jacobse, A. E., & Harskamp, E. (2011). A meta-Analysis of the Effects of instructional interventions on students' mathematics achievement. Groningen: GION, Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijksuniversiteit Groningen.

Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and special education*, 24(2), 97-114.

Other references

Askew, M., Brown, M., Rhodes, V., Johnson, D. C., & Wiliam, D. (1997). *Effective teachers of numeracy*. London: King's College.

Bloom, B. S. (1968, May). Mastery learning. In Evaluation comment (Vol. 1, No. 2). Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs. Borko, H., & Wildman, T. (1986). *Recent research on instruction. Beginning teacher assistance program*. Richmond, VA: Department of Education, Commonwealth of Virginia.

Brown, A. L., & Campione, J. C. (1990). Interactive learning environments and the teaching of science and mathematics. In M. Gardner et al. (Eds.), *Toward a scientific practice of science education*. Hillsdale, NJ: Erlbaum.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14.

McMullen, F., & Madelaine, A. (2014). Why is there so much resistance to Direct Instruction? *Australian Journal of Learning Difficulties*, 19(2), 137-151.

Rosenshine, B. (2008). Five meanings of direct instruction. *Center on Innovation & Improvement, Lincoln*.

Rosenshine, B. (2012). Principles of instruction: research-based strategies that all teachers should know. *American Educator*, 36(1), 12.

6.5 Mastery learning

What is the evidence regarding mastery learning in mathematics?

Evidence from US studies in the 1980s generally shows mastery approaches to be effective, particularly for mathematics attainment. However, very small effects were obtained when excluding all but the most rigorous studies carried out over longer time periods. Effects tend to be higher for primary rather than secondary learners and when programmes are teacher-paced, rather than student-paced. The US meta-analyses are focused on two structured mastery programmes, which are somewhat different from the kinds of mastery approaches currently being promoted in England. Only limited evidence is available on the latter, which suggests that, at best, the effects are small. There is a need for more research here.

Strength of evidence: MEDIUM

Findings

Bloom (1968) argued that when all learners in a class receive the same teaching, the learning achieved will vary considerably, whereas if instructional time and resources could be tailored to each learner's individual needs, a more uniform level of attainment could be achieved. He consequently advocated a *mastery* model of teaching in which teachers offered learners a variety of different approaches, with frequent feedback and extra time for those who struggled (which could take the form of tutoring, peer-assisted learning or extra homework). Content would be divided into small units, with tests at the end of each, and progression would be permitted only if learners exceeded a high threshold (such as 80%) on the tests. Alongside this would be enrichment tasks for those who had mastered the main ideas. Mastery has many similarities to direct instruction (see the module on explicit teaching), but differs in that in mastery, learners may be presented with alternative strategies.

In recent years in England, *mastery learning* has come to refer to a collection of practices used in high-performing jurisdictions, such as Shanghai and Singapore, which are focused on a coherent and consistent approach to using manipulatives and representations. In common with Bloom, mastery learning in this sense aims for a more uniform degree of learning and for all learners to achieve a deep understanding of and competence in the central ideas of a topic. However, this is through interactive whole-class teaching and common lesson content for all pupils (NCETM, 2016). This approach also encourages carefully sequenced lessons and early intervention to support learners who are struggling.

Fairly high to very high effect sizes are generally found for mastery approaches in mathematics (Guskey, & Pigott, 1988; Kulik, Kulik, & Bangert-Drowns, 1990; Rakes et al., 2010), particularly at primary (Guskey & Pigott, 1988), and particular where learners are forced to move through material at the teacher's pace, rather than at their own (Kulik, Kulik, & Bangert-Drowns, 1990). It also seems to be important for strong effects that learners are required to perform at a high level on unit tests (e.g., to obtain 80-100% correct) before proceeding, and that they receive feedback (Kulik, Kulik, & Bangert-Drowns, 1990). Low-attaining pupils may benefit more from mastery learning than high-attaining students (EEF, 2017).

In contrast to these findings, Slavin's (1987) best-evidence synthesis examined the results of seven studies which met his stringent criteria, which included longer interventions and the use of *standardised* achievement measures (rather than

experimenter-made measures). He found an overall ES of essentially zero (0.04), which suggests that caution should be exercised over the findings of the other meta-analyses. He argued that results in other meta-analyses could have been inflated by experimenter-designed instruments (i.e., teaching to the test) and effects deriving from increased instructional time and more frequent criterion-based feedback, rather than mastery *per se*. This raises an important issue. Mastery learning, like direct instruction, may be particularly effective in addressing specific topics or procedures, as might be measured by experimenter-designed instruments. Moreover, like direct instruction (McMullen, & Madelaine, 2014; Rosenshine, 2008), mastery claims to address conceptual as well as procedural knowledge. However, it is less clear that these approaches help learners to develop connections between areas of mathematics, or generic problem-solving skills, or the vital area of metacognition. It has been suggested that mastery learning may be most effective as an occasional or supplementary teaching approach; it appears that the impact of mastery learning decreases for programmes longer than around 12 weeks (EEF, 2017).

Evidence concerning the efficacy of the mastery learning approach currently being promoted in England is limited. Unlike the mastery programmes based on Bloom's work, key aspects of the approach such as early intervention and careful sequencing are not specified in detail. Rather, they are communicated through general principles (e.g., NCETM, 2016). It is left to schools and teachers to develop these principles into specific practices. The shift towards a mastery approach involves substantial professional change, and it seems unlikely that this will be achieved without considerable support, resources and professional development, such as that which was made available for the National Numeracy Strategy (Machin & McNally, 2009). However, one whole-school programme, *Mathematics Mastery*, provides a structure for schools that aims to deepen pupils' conceptual understanding of key mathematical concepts by covering fewer topics in more depth, emphasising problem solving and adopting the Concrete-Pictorial-Abstract approach commonly used in Singapore. Two RCTs of *Mathematics Mastery* carried out by the EEF (Jerrim & Vignoles, 2015), one at primary and the other at secondary, did not find effects that were significantly different from zero, but when these separate studies were combined a very small positive ES of 0.07 was produced. It is possible that the small sizes of the effects (if any) could be due to the fact that, unlike the US programmes, *Mathematics Mastery* does not wait to start new topics until a high level of proficiency has been achieved by all students on preceding material.

Evidence base

The evidence base is dated, but three meta-analyses (Guskey, & Pigott, 1988; Kulik, Kulik, & Bangert-Drowns, 1990; Rakes et al., 2010) report effect sizes for mastery in mathematics, while a fourth indicates an overall effect across subjects but where three of the seven studies synthesised are from mathematics.

Guskey and Pigott (1988) found a mathematics ES of 0.70, which was larger than for other subjects (0.50 for science and 0.60 for language arts). They also found that mastery had significantly higher effects for primary level.

In their systematic review of algebra instructional improvement strategies among older (Grades 9-college) students, Rakes et al. (2010) also found an overall ES of 0.469 for mastery in mathematics.

Kulik, Kulik and Bangert-Drowns (1990) investigated two different approaches to mastery – Bloom’s *Learning for Mastery* (Bloom, 1968), where all learners move through the material at the same pace, and Keller’s *Personalized System of Instruction* (Keller, 1968), where learners work through the lessons at their own pace. The authors found an overall ES of 0.47 for mathematics, which was similar to that for science but lower than that for social science, and no difference in ES between the two approaches. Eleven of the studies reported by Kulik, Kulik and Bangert-Drowns (1990) examined student performance on *delayed* post-tests, about 8 weeks after the intervention was concluded. The average ES obtained was 0.71, which was not significantly different from the average ES at the end of instruction across these same 11 studies, which was 0.60. Kulik, Kulik and Bangert-Drowns (1990) reported that ESs as large as 0.8 were “common” (p. 286) in studies which:

- focused on social sciences rather than on mathematics, the natural sciences, or humanities;
- used locally-developed rather than nationally standardised tests as measures of learner achievement;
- required learners to move through material at the *teacher’s* pace, rather than at individual students’;
- required students to perform at a high level on unit tests (e.g., obtain 100% correct);
- the control students received less test feedback than the intervention students did.

Directness

Most of the research synthesised is from the US, using dated programmes that were not designed for England. The exceptions to this are the two studies of *Mathematics Mastery*, which showed very small or no effects.

It may be that the level of prescription associated with some versions of mastery could be unattractive to mathematics teachers in England. Research is needed into the kinds of mastery approaches currently being advocated in England.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Most of the research is located in the US and, aside from <i>Mathematics Mastery</i> , the programmes were not designed for England and are different in approach from the mastery approaches currently being promoted in England.
How the intervention was defined and operationalised	2	Social validity: Teachers in England may find the high level of prescription in some kinds of mastery teaching unacceptable.
Any reasons for possible ES inflation	2	See Slavin’s (1987) critique.
Any focus on particular topic areas	2	Mastery may be more effective for teaching specific procedures and less effective for developing conceptual understanding, metacognition, connections and problem solving.
Age of participants	3	

Overview of Effects

Study	Effect size	No. of studies	Quality judgment (1 low to 3 high)	Notes
Guskey, & Pigott (1988)	0.70	36 maths	2	The mathematics effects were not homogenous, but split into topics and levels (algebra, geometry, probability, elementary, general high school). Only showed homogeneity for probability, so considerable variation is not explained.
Kulik, Kulik, & Bangert-Drowns (1990)	0.47	25 maths	2	Compared two approaches: Bloom's Learning for Mastery (LFM) and Keller's Personalized System of Instruction (PSI) and found no evidence of a difference between them.
Rakes et al. (2010)	0.469	4	3	Value obtained from supplementary data provided by the author. However, the ES is only based on studies with older students: Grades 9, 10 and college. All are pre-1987, but no overlap with Kulik et al. or Guskey, & Pigott.
Slavin (1987)	0.04	7 (3 maths)	2	Some key criticisms of mastery approaches. This is a best-evidence synthesis, so a bit more than a meta-analysis.

Mathematics Mastery Studies

(Note: these are two single studies, rather than a meta-analysis.)

Study	Effect size	No. of pupils	Notes
Jerrim & Vignoles (2015)	0.073	10,114 (in 127 schools)	For primary (4,176 pupils in 83 schools), the ES is 0.10 with 95% CI [-0.01, +0.21]; for secondary (5,938 pupils in 44 schools) the ES is 0.06 with 95% CI [-0.04 to +0.15], so both are non-significant. When combined in meta-analysis, the overall ES is 0.073 with

			95% CI [0.004 to 0.142], so <i>just</i> significant.
--	--	--	--

Second-order meta-analysis included

Education Endowment Foundation (2017). *Teaching & Learning Toolkit: Mastery learning*. London: EEF.

Meta-analyses

Guskey, T. R., & Pigott, T. D. (1988). Research on Group-Based Mastery Learning Programs: A Meta-Analysis. *The Journal of Educational Research*, 81(4), 197-216. doi:10.1080/00220671.1988.10885824

Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of Mastery Learning Programs: A Meta-Analysis. *Review of Educational Research*, 60(2), 265-299. doi:doi:10.3102/00346543060002265

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of Instructional Improvement in Algebra A Systematic Review and Meta-Analysis. *Review of Educational Research*, 80(3), 372-400.

Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, 57(2), 175-213. doi:10.3102/00346543057002175

Meta-analyses Excluded

Waxman, H. C., Wang, M. C., Anderson, K. A., & Walberg, H. J. (1985). Adaptive Education and Student Outcomes: A Quantitative Synthesis. *The Journal of Educational Research*, 78(4), 228-236. doi:10.1080/00220671.1985.10885607

[Subject differences in general are reported, but nothing specific for mathematics.]

Other references

Bloom, B. S. (1968, May). Mastery learning. In Evaluation comment (Vol. 1, No. 2). Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs.

Jerrim, J., & Vignoles (2015). *Mathematics Mastery: Overarching Summary Report*. Education Endowment Foundation.

Keller, F. S. (1968). "Good-bye, teacher..." *Journal of Applied Behavioral Analysis*, 1, 79-89.

Machin, S., & McNally, S. (2009). *The Three Rs: What Scope is There for Literacy and Numeracy Policies to Raise Pupil Achievement?*

McMullen, F., & Madelaine, A. (2014). Why is there so much resistance to Direct Instruction? *Australian Journal of Learning Difficulties*, 19(2), 137-151.

NCETM (2016). *The Essence of Maths Teaching for Mastery*. Retrieved from <https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf>

6.6 Problem solving

What is the evidence regarding problem solving, inquiry-based learning and related approaches in mathematics?

Inquiry-based learning (IBL) and similar approaches involve posing mathematical problems for learners to solve without teaching a solution method beforehand. Guided discovery can be more enjoyable and memorable than merely being told, and IBL has the potential to enable learners to develop generic mathematical skills, which are important for life and the workplace. However, mathematical exploration can exert a heavy cognitive load, which may interfere with efficient learning. Teachers need to scaffold learning and employ other approaches alongside IBL, including explicit teaching. Problem solving should be an integral part of the mathematics curriculum, and is appropriate for learners at all levels of attainment. Teachers need to choose problems carefully, and, in addition to more routine tasks, include problems for which learners do not have well-rehearsed, ready-made methods. Learners benefit from using and comparing different problem-solving strategies and methods and from being taught how to use visual representations when problem solving. Teachers should encourage learners to use worked examples to compare and analyse different approaches, and draw learners' attention to the underlying mathematical structure. Learners should be helped to monitor, reflect on and discuss the experience of solving the problem, so that solving the problem does not become an end in itself. At primary, it appears to be more important to focus on making sense of representing the problem, rather than on necessarily solving it.

Strength of evidence (IBL): LOW

Strength of evidence (use of problem solving): MEDIUM

Introduction

Problem solving is crucial to the use and application of mathematics in the world beyond school (e.g., Hodgen & Marks, 2013; see also ACME, 2011, 2016). As a result, problem-solving skills are an important aim of school mathematics education as set out in the National Curriculum for England. However, problem solving encompasses a range of tasks. At one extreme, any task presented to a student may be defined as 'a problem', including, in much of the US literature, 'word problems', which are often direct applications of a given method in a real-world context. At the other extreme, problem solving may be understood to take place only when students are presented with a task for which they have no immediately applicable method, and consequently have to devise and pursue their own approach.

Problem solving and inquiry provoke heated debate concerning how best they should be taught and the extent to which learners should master the 'basics' of mathematics first. Nevertheless, as noted in the overview to this document, the literature on learners' development suggests that problem solving is needed for learners to develop generic mathematical skills. In this module, we examine the evidence relating to these issues and the role of problem solving, inquiry-based learning and related approaches in mathematics learning more widely.

Findings

We found nine meta-analyses relevant to problem solving (11 originally but two were excluded). In addition, we identified one US-focused *What Works Clearinghouse* (WWC) practitioner guide on the teaching of problem solving. The meta-analyses address different, but related, constructs, and, in particular, define problem solving in very different ways. Eight of the 11 meta-analyses were concerned with approaches to teaching, such as inquiry-based learning, problem-based learning, the teaching of heuristics, (guided) discovery learning and integrative approaches. The remaining three meta-analyses, and the WWC practitioner guide, addressed the use of problems and the teaching of problem solving more directly. Hence, we present our findings under these two categories: the effects of inquiry-based learning and related approaches to teaching, and the use and teaching of problem solving.

The effects of inquiry-based learning and related approaches to teaching

Inquiry-based learning (IBL) and *problem-based learning* are *active learning*, *student-centred* teaching approaches in which students are presented with a scenario and encouraged to specify their own questions, locate the resources they need to answer them, and *investigate* the situation, so as to arrive at a solution. Problems may be located in the real world (i.e., *modelling problems*) or set in the context of pure mathematics. IBL approaches tend to rely on the use of collaborative learning (see module on collaborative learning) and it is argued that IBL trains learners in skills (such as communication) that are important for life and the workplace. It is also argued that *discovering* information may be more enjoyable and memorable than merely receiving it passively (Hmelo-Silver, Duncan, & Chinn, 2007).

However, it has also been strongly argued that approaches involving minimal guidance are less effective than *explicit teaching* (see module on explicit teaching) because they fail to allow for learners' limited working memory and expect novice learners to behave like experts, even though they do not have the necessary bank of knowledge to do this (Kirschner, Sweller, & Clark, 2006). Kirschner, Sweller and Clark (2006) argue that the "way an expert works in his or her domain ... is not equivalent to the way one learns in that area" (p. 78), and thus "teaching of a discipline as inquiry" should not be confused with "teaching of a discipline *by* inquiry" (p. 78, emphasis added). Exploration of a complex environment generates a heavy cognitive load that may be detrimental to learning. This is less of a problem for more knowledgeable "expert" learners, but disproportionately disadvantages low-attaining learners; although they may enjoy IBL approaches more, they learn less (p. 82).

In response to this, it has been countered that IBL approaches are not in fact minimally guided, as portrayed, and employ extensive scaffolding, which reduces cognitive load (Hmelo-Silver, Duncan, & Chinn, 2007). It may also be that cognitive load may be well managed if *worked examples* are used, where learners can be invited to reflect on the strategy and tactics of solving the problem, rather than the details of the calculations. This has similarities to the *neriage* phase of the Japanese problem-solving lesson, in which "the lesson begins when the problem is solved" (Takahashi, 2016; see also the module on metacognition and thinking skills). Solving

the problem must not become an end in itself, if the goal is to learn about how to solve future (as yet unknown) problems. Teaching problem solving is effective where learners are able to transfer their knowledge to different applications. It is known that learner disaffection is a huge problem, particularly at key stage 3 (e.g., Nardi et al., 2003; Brown et al., 2008), and Savelsbergh et al. (2016) provide evidence to suggest that innovative IBL approaches can have a positive effect on attitudes, with a neutral or positive effect on attainment.

Scheerens et al. (2007) examined school and teaching effectiveness using a wide range of studies including observational/correlational studies. In particular, they looked at constructivist-oriented learning strategies (*constructivist teaching* is a term commonly used in the US in the 1980s & 1990s and is broadly similar to student-centred teaching [Simon, 1995]). They compared this to structured, direct teaching and teacher-orchestrated classroom management, finding similar, small ESs of around 0.1 for all of these. Scheerens et al. commented that:

effective teaching is a matter of clear structuring and challenging presentation and a supportive climate and meta-cognitive training. The results indicate that these main orientations to teaching are all important, and that effective teaching is not dependent on a singular strategy or approach. (p. 131)

This suggests that in ordinary, non-experimental classrooms, the differences on attainment between IBL and teacher-centred approaches may not be very pronounced, and a judicious balance may be optimal.

Preston (2007) found that student achievement was higher with *student-centered instruction*, in which students *actively* participated in discussion, than with teacher-centered instruction, where the teacher did most of the talking (ESs around 0.54).

Becker and Park (2011) found that *integrative approaches* showed larger ESs at primary than at the college level, and the integration of all four parts of “STEM” gave the largest effect size (0.63).

Gersten et al. (2009) defined a *heuristic* as “a method or strategy that exemplifies a generic approach for solving a problem” (p. 1210). As an example, they suggest the following generic approach: “Read the problem. Highlight the key words. Solve the problem. Check your work.” Heuristics are not problem-specific and can be applied to different types of problems, and may involve more structured approaches to analysing and representing a problem. Gersten et al. (2009) found a huge ES of 1.56 for teaching heuristics (compared with 1.22 for explicit instruction). These very high ESs are probably inflated because the meta-analysis focused on learners with learning disabilities; however, it may be fair to conclude that these findings suggest that heuristics could be comparable with explicit teaching in terms of its capacity to raise attainment. Explicit teaching and heuristics may be complementary approaches, explicit teaching being particularly appropriate for important techniques that learners will need to use again and again, and heuristics being vital to help learners develop flexibility and the ability to tackle the unknown.

Finally, in a study from the 1980s, Athappilly, Smidchens and Kofel (1983) found small ESs in favour of “modern mathematics” (focused on abstract, early-20th century

mathematics) relative to traditional mathematics (attainment, 0.24; attitude, 0.12), although we observe that this speaks to a rather dated debate.

The use and teaching of problem-solving

As stated above, mathematical problem solving takes place when a learner tackles a task for which they do not have a suitable readily-available solution method (NCTM, 2000). In practice, this means that a classroom task could be regarded as a “problem” if the teacher has not, immediately prior to the task, taught an explicit method for solving it. Typically, guidance on problem solving recommends the use of a wide range of problem types (e.g., NCTM, 2000; Woodward et al., 2012). However, much of the research literature focuses on *word problems*. Of the 487 studies included in Hembree’s (1992) meta-analysis, the vast majority focused on standard, or *routine*, word or story problems, which require the solver to translate the story into a mathematical calculation, and relatively few examined non-standard problems, for which the solver does not have a well-rehearsed and ready-made method. Only one study focused on real-world problems and none examined a problem type which Hembree terms ‘puzzles’, which require unusual or creative strategies.

Hembree (1992) provided evidence of the efficacy of problem solving ($ES = 0.77$ relative to no problems), and found that problem solving is appropriate for students at all attainment levels. However, there is considerable variation. There is some evidence of a positive impact on students' performance for problem solving with instruction over no instruction. Hembree also reported benefits resulting from teachers trained in heuristics. From Grade 6 onwards, heuristics training appeared to give increasing improvements in problem-solving performance. For example, “[i]nstruction in diagram drawing and translation from words to mathematics also offer large effects toward better performance. Explicit training appears essential; these subskills do not appear to derive from practice without direction and oversight” (p. 267). He also indicated a strong effect for training learners to represent problems ($d=1.16$), and that physical manipulatives help students to do this. There is some evidence to suggest that primary learners may benefit more from representing problems than from necessarily solving them or being taught problem-solving heuristics. Hembree also found that reading ability does not appear to be a critical requirement for problem solving.

Rosli et al. (2014) found varied ES s for problem *posing*, with some evidence of effects on knowledge as well as skills, concluding that problem-posing “activities provide considerable benefits for: mathematics achievement, problem solving skills, levels of problems posed, and attitudes toward mathematics”.

Sokolowski (2015) explored whether mathematical modelling helps students to understand and apply mathematics concepts. They found 13 studies with an ES of 0.69 and advocated a wider implementation of modelling in school. However, some of their ES s are likely to be inflated. Teacher effects are likely to be very strong.

As already noted, we did identify a *What Works Clearinghouse* (WWC) practitioner guide on “*Improving mathematical problem solving in grades 4 through 8*”

(Woodward et al., 2012). Their recommendations in relation to problem solving include:

1. **Prepare and use them in whole-class instruction:** The WWC panel recommended that problem-solving should be an integral part of the mathematics curriculum and that teachers should deliberately choose a variety of problems, including both routine (standard) and non-routine (non-standard) problems, and considering learners' mathematical knowledge. When selecting problems and planning teaching, teachers should consider issues relating to context or language in order to enable learners to understand a problem.
2. **Assist students in monitoring and reflecting on the problem-solving process:** Learners solve mathematical problems better when they regulate their thinking through monitoring and reflecting (see metacognition module). The panel identified three evidence-based effective approaches: (i) providing prompts to encourage learners to monitor and reflect during problem solving, (ii) teachers modelling how to monitor and reflect during problem solving, and (iii) using and building upon learners' ideas.
3. **Teach students how to use visual representations:** The panel identified three evidence-based effective approaches: (i) teachers should deliberately select visual representations that are appropriate to the problem and for the learners, (ii) the use of think-aloud and discussion to teach learners how to represent problems, and (iii) demonstrating how to translate visual representations into mathematical notation and statements (see manipulatives and representations module).
4. **Expose students to multiple problem-solving strategies:** The panel identified three evidence-based effective approaches: (i) teach learners different problem-solving strategies, (ii) use worked examples to enable learners to compare different strategies, and (iii) encourage learners to generate and share different problem-solving strategies.
5. **Help students recognise and articulate mathematical concepts and notation:** The panel identified three evidence-based effective approaches: (i) highlight and describe relevant mathematical ideas and concepts used by learners during problem-solving, (ii) ask learners to explain the steps in worked examples and explain why they work, and (iii) help learners to understand algebraic notation (see Algebra section of mathematical topics module).

Atkinson et al. (2000) advocate using, for each type of problem, multiple examples, where the surface features vary from example to example in order to draw attention to a consistent, deeper structure. They stress the *active* use of worked examples by suggesting that learners be required to actively self-explain the solutions, and they point out that worked examples are particularly beneficial at the early stages of skill development.

Evidence base

None of the meta-analyses here are of the highest methodological quality, and the most relevant one (Hembree, 1992) is dated.

The WWC practitioner guidance judged the evidence to be strong for two recommendations (monitoring and reflecting, and using visual representations), to be moderate for two recommendations (multiple strategies and recognition/articulation of mathematical concepts and notation), and to be minimal for one recommendation (the preparation and use of problems).

There is a pressing need for an up-to-date meta-analysis looking at problem solving. There is also a great need for researchers to develop standardised tests that assess problem solving, as using specific researcher-designed tests tends to inflate ESs.

We draw on Gersten (2009) only tangentially, as it is focused on learners with learning disabilities, which is likely to inflate ESs. The ESs used are based on small sets of studies ($k = 4$ for heuristics and $k = 11$ for explicit teaching) and the Q statistic is high, meaning that all the variation is not explained. This suggests that the efficacy of both explicit teaching and heuristic strategies may be dependent on other factors, such as the mathematical topic or context.

Sokolowski (2015) looked at studies in the high school and college age, and the vast majority of measures used were researcher-designed, which may have inflated the ESs reported.

Directness

Our overall judgment is that the findings of the meta-analyses have moderate directness. Despite differences in the US and English curricula, the WWC Practice Guide (Woodward et al., 2012) is judged to highlight approaches that would be applicable in the English context.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Studies mostly carried out in the US, where the teaching culture is somewhat different from England. However, a general absence of IBL teaching is a feature of both countries.
How the intervention was defined and operationalised	2	Problems of varying definitions quite serious.
Any reasons for possible ES inflation	1	Some studies report for learners with learning disabilities, which inflates ESs. Frequently researcher-designed tests, which also inflate ESs.
Any focus on particular topic areas	3	
Age of participants	3	Mostly OK.

Overview of effects

Study	Effect size (<i>d</i>)	No. of studies	Quality judgment (1 low to 3 high)	Notes
Savelsbergh et al. (2016)	attitude: 0.35 attainment: 0.78	61 40	3	Examined the effects of innovative science and mathematics teaching on student attitudes and achievement.
Scheerens et al. (2007)	0.09 (structured, direct, mastery,...); 0.14 (constructivist-oriented ...)	165 (structured, direct, mastery,...); 542(constructivist-oriented ...)	2	Review and meta-analyses of school and teaching effectiveness.
Preston (2007)	0.56 (primary); 0.52 (secondary)	18	2	Examined student-centered versus teacher-centered mathematics instruction.
Athappilly, Smidchens, & Kofel (1983)	0.24 (achievement), 0.12 (attitude), both in favor of modern	660 (attainment), 150 (attitude)	2	Very dated study which compared modern mathematics with traditional mathematics.
<i>Integrative approaches</i>				
Becker & Park (2011)	0.63	28	1	Examined the impact of interventions aimed at the integration of science, technology, engineering,

				and mathematics disciplines.
Scheerens et al. (2007)	0.09	90	2	Review and meta-analyses of school and teaching effectiveness.
<i>Problem solving</i>				
Hembree (1992)	0.77	487	2	Dated study looked at learning and teaching of problem solving.
Rosli et al. (2014)	0.76 – 1.31	14	2	Looked at problem-posing activities.
Sokolowski (2015)	0.69	13	2	Looked at effects of Mathematical Modelling on Students' Achievement.

Systematic review on the teaching of problem-solving	No of studies (k)	Comment
Siegler et al. (2010) (WWC Practice Guidance): <i>Improving mathematical problem solving in grades 4 through 8</i>	-	Uses <i>What Work Clearinghouse</i> standards
Prepare problems and use them in whole-class instruction	6	Minimal evidence base
Assist students in monitoring and reflecting on the problem-solving process	12	Strong evidence base
Teach students how to use visual representations	7	Strong evidence base
Expose students to multiple problem-solving strategies	14	Moderate evidence base
Help students recognize and articulate mathematical concepts and notation	6	Moderate evidence base

References

Meta-analyses

- Athappilly, K., Smidchens, U., & Kofel, J. W. (1983). A computer-based meta-analysis of the effects of modern mathematics in comparison with traditional mathematics. *Educational Evaluation and Policy Analysis*, 485-493.
- Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education: Innovations and Research*, 12(5/6), 23.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202-1242
- Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. *Journal for Research in Mathematics Education*, 242-273.
- Preston, J. A. (2007). Student-centered versus teacher-centered mathematics instruction: A meta-analysis. Doctoral Thesis, Indiana University of Pennsylvania. ProQuest UMI 3289778.
- Rosli, R., Capraro, M. M., & Capraro, R. M. (2014). The effects of problem posing on student mathematical learning: A meta-analysis. *International Education Studies*, 7(13), 227
- Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, 19, 158-172.
- Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). Review and meta-analyses of school and teaching effectiveness. *Enschede: Department of Educational Organisation and Management, University of Twente*.
- Sokolowski, A. (2015). The Effects of Mathematical Modelling on Students' Achievement-Meta-Analysis of Research. *IAFOR Journal of Education*, 3(1), 93-114.

Meta-analyses Excluded

- 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(4), 207-225. [Superseded by Zhang & Xin (2012).]
- Zhang, D., & Xin, Y. P. (2012). A follow-up meta-analysis for word-problem-solving interventions for students with mathematics difficulties. *The Journal of educational research*, 105(5), 303-318. [Focus on learning difficulties.]

Systematic review

- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research.

Review of Educational Research, 70(2), 181-214.
doi:10.3102/00346543070002181

Woodward, J., Beckmann, S., Driscoll, M., Franke, M. L., Herzig, P., Jitendra, A., Koedinger, K. R., Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide (NCEE 2012-4055)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Other references

Advisory Committee on Mathematics Education [ACME]. (2011). *Mathematical Needs: Mathematics in the workplace and in Higher Education*. London: Royal Society.

Advisory Committee on Mathematics Education [ACME]. (2016). *Problem solving in mathematics: realising the vision through better assessment*. London: Royal Society.

Brown, M., Brown, P., & Bibby, T. (2008). "I would rather die": Attitudes of 16 year-olds towards their future participation in mathematics. *Research in Mathematics Education*, 10(1), 3-18.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational psychologist*, 42(2), 99-107.

Hodgen, J., & Marks, R. (2013). *The Employment Equation: Why our young people need more maths for today's jobs*. London: The Sutton Trust.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.

Nardi, E., & Steward, S. (2003). Is mathematics T.I.R.E.D? A profile of quiet disaffection in the secondary mathematics classroom. *British Educational Research Journal*, 29(3), 345-367.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26, 114

Takahashi, A. (2016). Recent Trends in Japanese Mathematics Textbooks for Elementary Grades: Supporting Teachers to Teach Mathematics through Problem Solving. *Universal Journal of Educational Research*, 4(2), 313-319.

6.7 Peer and cross-age tutoring

What are the effects of using peer and cross-age tutoring on the learning of mathematics?

Peer and cross-age tutoring appear to be beneficial for tutors, tutees and teachers and involve little monetary cost, potentially freeing up the teacher to implement other interventions. Cross-age tutoring returns higher effects, but is based on more limited evidence. Peer-tutoring effects are variable, but are not negative. Caution should be taken when implementing tutoring approaches with learners with learning difficulties.

Strength of evidence: MEDIUM

Definitions

Cross-age tutoring involves an older learner (in a higher year) working with a younger learner, whereas peer-tutoring involves two learners of the same age working together, one in the role of tutor, the other as tutee. Gersten (2009) noted in his review that, although studies of peer-tutoring date back 50 years, it is still often regarded as a relatively novel approach.

Findings

We found no meta-analyses that examined peer or cross-age tutoring exclusively in the context of mathematics. Within meta-analyses considering a broad range of instructional interventions in mathematics, cross-age and peer-tutoring were considered in two analyses and peer-tutoring solely in a further seven. Five of these analyses focused on interventions for low-attaining or SEND learners. We also include one meta-analysis looking at peer-tutoring in general, with mathematics as a moderator, so we draw on 10 meta-analyses of cross/peer-tutoring in total.

The pooled ES for cross-age tutoring on general learners in mathematics (as tutees) is 0.79 (Hartley, 1977). For learners with LD this rises to 1.02 (Gersten et al., 2009), although this result should be interpreted with caution, as it is based on only two studies and a restricted range of learners. Cross-age tutoring has been repeatedly reported as the most effective form of tutoring, but may be difficult to organise. Training learners as tutors improves the effectiveness of tutoring interventions, but effectiveness can vary, particularly with EAL, SEN and low-attaining learners (Lloyd et al., 2015).

Pooled ESs for peer-tutoring on general learners in mathematics (as tutees) range from 0.27 to 0.60. Where moderator analyses were conducted, results are either significantly higher for mathematics or show no significant difference between mathematics and other subjects. In one meta-analysis (Leung, 2015), a greater range of subjects was examined – physical education, arts, science and technology and psychology – and it appears that these subjects return higher ESs than do mathematics and reading, although there were many more studies of mathematics and reading.

Two meta-analyses report similar ESs on general learners as tutors of 0.58 and 0.62 (Hartley, 1977; Cohen, Kulik and Kulik, 1982). For low-attaining learners, ESs of peer-tutoring (tutee and tutor combined) are 0.66 and 0.76 (Baker, Gersten & Lee, 2002; Lee, 2000). However, for LD learners ESs vary considerably from -0.09 (Kroesbergen & Van Luit, 2003) – although this should be treated with significant caution due to a range of methodological factors – to 0.76 (Lee, 2000).

Overall, the meta-analyses suggest that although peer-tutoring results are variable, the approach is not damaging for the general population or low-attaining learners, with all reported ESs being positive. Caution should be taken in implementing peer-tutoring with very weak LD learners, who may struggle with any form of peer-collaborative working and may reap more benefit from cross-age tutoring. Tutors require training and support, and tutoring situations need structure (Baker et al., 2002; Gersten, 2009). Lloyd et al.'s (2015) review notes that the tutor (learner) training for 'Shared Maths' focussed on how to understand and respond to mathematical questions (as opposed to general tutor training), and it may be that a mathematical focus to the training is important. Kroesbergen & Van Luit (2003) found the effects of peer-tutoring to be less than those of other interventions, which they suggest may be due to peers being less capable than teachers of perceiving other learners' mathematical needs. Baker et al. (2002), Hartley (1977) and Othman (1996) all conclude that peer-tutoring is beneficial for tutors (who develop responsibility and a deeper understanding of the material), tutees (who are less reluctant to ask questions of a peer) and teachers (who are freed up for other tasks), and involves little monetary cost (WSIPP, 2017).

Evidence base

Our findings come from eight meta-analyses which address peer-tutoring and two which address both cross-age tutoring and peer-tutoring. In every case, these findings are sub-sets of a wider meta-analysis. The two meta-analyses examining cross-age tutoring synthesised a total of 32 studies (note that Gersten et al. [2009] included only two of these 32 studies) and covered the date period 1962 to 2003. The 10 meta-analyses examining peer-tutoring synthesised a total of 299 studies (including studies outside of mathematics) and covered the date period 1961 to 2012. As discussed above, the ESs across the studies for peer-tutoring show some variability and a lot of the variation is not understood. The included meta-analyses predominantly have medium or high quality ratings. Although there is a fairly high degree of overlap in the included studies within each full meta-analysis (ranging from 39% to 68% for all post-2000 meta-analyses), the authors do not provide the information needed to ascertain the degree of overlap in included studies related to peer-tutoring, and there appear to be few robust studies in this area.

With regard to the comparison between mathematics and other subjects, Leung's (2015) meta-analysis synthesised far fewer studies in physical education, arts, science and technology and psychology than in mathematics (k=3 to k=6 compared with 20 studies in mathematics and 31 in reading).

Meta-analysis	k (for tutoring)	Quality	Date Range
Baker, Gersten, & Lee (2002)	6	2	1982-1999
Chen (2004)	5	3	1977-2003
Cohen, Kulik & Kulik (1982)	65 overall (11 for mathematics)	2	1961-1980
Gersten et al. (2009)	2 (cross-age) 6 (peer)	3	1982-2003

Hartley (1977)	29 (cross-age and peer combined)	2	1962-1976
Kroesbergen & Van Luit (2003)	10	2	1985-2000
Lee (2000)	10	2	1971-1998
Leung (2015)	72 (20 for mathematics)	3	pre-2012
Othman (1996)	18	1	1970-1992
Rohrbeck et al. (2003)	90 overall (25 for mathematics)	3	1974-2000

Directness

In contrast to other forms of collaborative learning, cross-age and (particularly) peer-tutoring interventions are not in the main delivered through particular structured programmes.

These findings are based on studies which are predominantly located in the US. Despite cultural differences, we judge that the findings may be transferable to the English context. The variation in the effects suggests that the implementation of peer-tutoring may be crucial to its efficacy. One recent trial at primary mathematics in England showed no effect for a cross-age tutoring intervention (Lloyd et al., 2015).

Where the meta-analyses reviewed in the Education Endowment Foundation toolkit (Higgins et al., 2013) focus on mathematics and meet our inclusion criteria, we have included them here. Higgins et al. (2013) report a range of ESs for peer-tutoring in general ($d = 0.35$ to $d = 0.59$, based on five meta-analyses published between 1982 and 2014),² for the effects of peer-tutoring on tutors and tutees ($d = 0.33$ & 0.65 and to $d = 0.40$ & 0.59 , respectively, based on two meta-analyses published in 1982 and 1985), and cross-age tutoring ($d = 1.05$, based on one meta-analysis published in 2010). Given the extent of this evidence base and the need to understand implementation better, there may be some value in synthesising the results of these meta-analyses, in particular to identify potential factors that may aid or hinder the effective implementation of peer-tutoring.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Studies mostly carried out in the US, where the teaching culture is somewhat different from England.
How the intervention was defined and operationalised	3	
Any reasons for possible ES inflation	3	

² One of the meta-analyses was based on single-subject designs and is not reported here.

Any focus on particular topic areas	3	
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Learners in general</i>			
<i>Effect of cross-age tutoring on tutees</i>			
Hartley (1977)	0.79	29	No CIs given k=29 is for all types of tutoring combined; breakdown of number of studies for cross/peer and tutor/tutee not given.
<i>Effect of peer-tutoring on tutees</i>			
Cohen, Kulik & Kulik (1982)	0.60	18	No CIs given Reading ES=0.29 (k=30) Other subjects ES=0.30 (k=4)
Hartley (1977)	0.52	17 effect sizes	No CIs given k=29 for all types of tutoring combined; breakdown of number of studies for cross/peer and tutor/tutee not given.
Leung (2015)	0.34 [0.27, 0.41]	20	Overall ES= 0.37 [0.29, 0.45] for the mixed effects model Reading ES=0.34 [0.31, 0.38] (k=31) N.B. while ESs for maths and reading are similar, there is a significant degree of unexplained variation. Other subjects (all with small k): Language ES= 0.15 [0.05, 0.25] (k=6) Science & technology ES= 0.45 [0.37, 0.53] (k=6) Physical Education ES= 0.90 [0.72, 1.07] (k=4) Arts ES= 0.82 [0.73, 0.91] (k=4)
Othman (1996)	0.30	18	
Rohrbeck et al. (2003)	0.27 [0.19, 0.34]	25	Overall ES=0.33 [0.29, 0.37], Reading ES=0.26 [0.19, 0.33] (k=19), the authors conclude that no significant differences in ES were found among PAL interventions implemented in mathematics and reading.

<i>Effect of peer-tutoring on tutors</i>			
Cohen, Kulik & Kulik (1982)	0.62	11	No CIs given Reading ES=0.21 (k=24)
<i>Effect of tutoring (peer and cross-age combined) on tutors</i>			
Hartley (1977)	0.58	18 effect sizes	No CIs given. ES overall for tutoring (peer and cross-age combined) on tutees was 0.63 k=29 is for all types of tutoring combined; breakdown of number of studies for cross/peer and tutor/tutee not given.
Low attaining learners			
<i>Effect of peer-tutoring on low attaining learner achievement (tutor and tutee combined)</i>			
Baker, Gersten, & Lee (2002)	0.66 [.42, .89]	6	The magnitude of effect sizes was greater on computation than general maths ability. The average effect size on computation problems was .62 (weighted), which was significantly greater than zero. On general maths achievement, the two effect sizes were .06 and .40, producing a weighted mean of .29 that was not significantly different from 0.
Lee (2000)	0.76 [0.19, 1.34]	6	
Learners with learning disabilities or special educational needs			
<i>Effect of LD cross-age tutoring on tutees</i>			
Gersten et al. (2009)	1.02 [0.57, 1.47]	2	
<i>Effect of LD peer tutoring on tutees</i>			
Chen (2004)	0.56	5	Results for group-design studies. Minimum ES=0.39, Maximum ES =1.47 No CIs reported.
Gersten et al. (2009)	0.14 [- 0.09, 0.32]	6	
Kroesbergen & Van Luit (2003)	-0.09	10 peer tutoring; 51 in control	The reported effect size of 0.87 is compared to a "constructed" control group effect of 0.96. This constructed control consists of the controls for all non-peer tutoring interventions combined.

			This constructed control group may then represent “business as usual”. Kroesbergen & Van Luit concludes that peer-tutoring has no effect. The meta-analysis aggregates experimental (with and without pre-tests) and single cases, therefore should be treated with caution.
<i>Effect of peer-tutoring on LD learner achievement (tutor and tutee combined)</i>			
Lee (2000)	0.76 [-1.10, 2.62]	4	
Effect of Peer-Assisted Learning Strategies peer-tutoring program on general (K-6) learners			
U.S. Department of Education, IES WWC (2013)	Average improvement index of 2 [range: -1 to 6]	1	WWC report (i.e. not a full meta-analysis) with only one study which met the WWC methodological and reporting standards. Study found no discernible effect on mathematics achievement.

References

Meta-analyses included

- Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, 103(1), 51-73.
- Chen, H. (2004). The efficacy of mathematics interventions for students with learning disabilities: A meta-analysis. Unpublished PhD, The University of Iowa.
- Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal*, 19(2), 237-248.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202-1242.
- Hartley, S.S. (1977) Meta-Analysis of the Effects of Individually Paced Instruction In Mathematics. Doctoral dissertation University of Colorado at Boulder.
- Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and special education*, 24(2), 97-114.
- Lee, D. S. (2000). A meta-analysis of mathematics interventions reported for 1971-1998 on the mathematics achievement of students identified with learning disabilities and students identified as low achieving. Doctoral Thesis, University of Oregon ProQuest UMI 9963449

- Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, 107(2), 558-579.
- Othman, N. (1996). The effects of cooperative learning and traditional mathematics instruction in grade K-12: A meta-analysis of findings. Doctoral Thesis, West Virginia University. ProQuest UMI 9716375
- Rohrbeck, C., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240-257.

Meta-analyses excluded

- Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39-55. [Excluded due to focus on single-case designs.]

Systematic reviews included

- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013). *Elementary School Mathematics intervention report: Peer-Assisted Learning Strategies*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Other references

- Dugan, J.J. (2007). A systematic review of interventions in secondary mathematics with at-risk students: Mapping the Literature Doctoral dissertation Colorado State University, Fort Collins Colorado. ProQuest UMI 3266389
- Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L. E., & Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*. London: Education Endowment Foundation.
- Lloyd, C., Edovald, T., Morris, S., Kiss, Z., Skipp, A., & Haywood, S. (2015) *Durham Shared Maths Project: Evaluation Report and Executive Summary*. London: EEF
- WSIPP (2017) Washington State Institute for Public Policy Benefit-Cost Results: Tutoring By Peers Pre-K to 12 Education.

6.8 Misconceptions

What is the evidence regarding misconceptions in mathematics?

Students' misconceptions arise naturally over time as a result of their attempts to make sense of their growing mathematical experience. Generally, misconceptions are the result of over-generalisation from within a restricted range of situations. Misconceptions should be viewed positively as evidence of students' sense making. Rather than confronting misconceptions in an attempt to expunge them, exploration and discussion can reveal to students the limits of applicability associated with the misconception, leading to more powerful and extendable conceptions that will aid students' subsequent mathematical development.

Strength of evidence: MEDIUM

Findings

A misconception is “a student conception that produces a systematic pattern of errors” (Smith, diSessa, & Roschelle, 1994, p. 119) and leads to perspectives that are not in harmony with accepted mathematical understanding. Much research has documented common misconceptions and misunderstandings which students develop in different mathematics topics.

Misconceptions arise out of students' prior learning, either from within the classroom or from the wider world. When viewed from the perspective of the students' previous experience, misconceptions make sense, because they explain some set of phenomena within a restricted range of contexts:

Most, if not all, commonly reported misconceptions represent knowledge that is functional but has been extended beyond its productive range of application. Misconceptions that are persistent and resistant to change are likely to have especially broad and strong experiential foundations. (Smith, diSessa, & Roschelle, 1994, p. 152)

For example, the “multiplication makes bigger, division makes smaller” conception is an accurate generalisation for numbers greater than 1. It is only when extended beyond this set of numbers that this conception becomes a *misconception*.

Misconceptions create problems for students when they lead to errors in calculation or reasoning. Typically, they are benign for a time, but, as subsequent mathematical concepts appear and have to be taken account of (e.g., numbers less than or equal to 1), they become problematic. Teachers need to take students' misconceptions seriously, and not dismiss them as nonsensical, by thinking about what prior experiences could have led to the students' particular misconceptions. As Smith, diSessa and Roschelle (1994, p. 124) put it, “misconceptions, especially those that are most robust, have their roots in productive and effective knowledge”, and this is why they can be quite stable, widespread and resistant to change.

It is often assumed that misconceptions must be uncovered and then confronted, so as to “overcome” them and replace them with correct concepts. Through *cognitive conflict*, the disparity between mathematical reality and what the student believes will become explicit, and then students will modify their beliefs accordingly. However, this is sometimes not effective. Students will often actively defend their misconceptions, and teaching that simply confronts students with evidence that they are wrong is thought by Smith, diSessa and Roschelle (1994, p. 153) to be

“misguided and unlikely to succeed”. Instead, it is necessary to explore how the misconception has arisen, the “partial truth” that it is built on, when it is valid and when and why it is not, in order to assist students, over a period of time, to generalise more substantially, so as to arrive at different and more useful conceptions of mathematics.

Evidence base

Smith, diSessa and Roschelle (1994) in their classic paper “Misconceptions reconceived” summarised knowledge about misconceptions and interpreted this from a constructivist perspective.

Many have catalogued and summarised students’ specific mathematical misconceptions in detail (e.g., Hart et al., 1981; Ryan & Williams, 2007). Reynolds and Muijs (1999) discussed awareness of misconceptions in the context of effective teaching of mathematics.

Directness

We have no concerns over the directness of these findings.

References

- Hart, K. M., Brown, M. L., Kuchemann, D. E., Kerslake, D., Ruddock, G., & McCartney, M. (1981). *Children’s understanding of mathematics: 11-16*. London: John Murray.
- Reynolds, D., & Muijs, D. (1999). The effective teaching of mathematics: A review of research. *School Leadership & Management*, 19(3), 273-288.
- Ryan, J., & Williams, J. (2007). *Children’s mathematics 4-15: Learning from errors and misconceptions*. McGraw-Hill Education.
- Smith III, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115-163.

6.9 Thinking skills, metacognition and self-regulation

To what extent does teaching thinking skills, metacognition and/or self-regulation improve mathematics learning?

Teaching thinking skills, metacognition and self-regulation can be effective in mathematics. However, there is a great deal of variation across studies. Implementing these approaches is not straightforward. The development of thinking skills, metacognition and self-regulation takes time (more so than other concepts), the duration of the intervention matters, and the role of the teacher is important. One thinking skills programme developed in England, Cognitive Acceleration in Mathematics Education (CAME), appears to be particularly promising. Strategies that encourage self-explanation and elaboration appear to be beneficial. There is some evidence to suggest that, in primary, focusing on cognitive strategies may be more effective, whereas, in secondary, focusing on learner motivation may be more important. Working memory and other aspects of executive function are associated with mathematical attainment, although there is no clear evidence for a causal relationship. A great deal of research has focused on ways of improving working memory. However, whilst working memory training improves performance on tests of working memory, it does not have an effect on mathematical attainment.

Strength of evidence: MEDIUM

Definitions

This question addresses one of the key aspects of the development of mathematical competency as discussed in Section 3. Metacognition is broadly defined as ‘thinking about thinking’ and the understanding of one’s thinking and learning processes. Self-regulation is related to metacognition and is defined as the dispositions (such as resilience, perseverance and motivation) to put one’s cognitive and metacognitive processes into practice. Cognitive strategies include aspects such as organisational skills, serving as pre-requisites for later metacognitive processes.

Thinking skills is a looser but related notion. Thinking skills interventions can be defined as approaches or programmes that are designed to develop learners’ cognitive, metacognitive and self-regulative knowledge and skills. Typically, thinking skills programmes focus either on generic thinking skills or on developing thinking skills in the context of a particular curriculum area, such as mathematics.

Executive function is “the set of cognitive skills required to direct behavior toward the attainment of a goal” (Jacob & Parkinson, 2015, p. 512). Working memory (WM) is commonly thought of as a subcomponent of executive function. It involves the brain’s “temporary storage” while engaging in “complex cognitive tasks” (Melby-Lervåg & Hulme, 2013, p. 270). A number of models and components of WM have been proposed.

Findings

Teaching thinking skills, metacognition and self-regulation can be effective in mathematics. We found a large number of recent meta-analyses in this area with a wide range of effects, some very large. For thinking skills, metacognitive and self-regulative interventions aimed at increasing attainment in mathematics – or aspects of mathematics – we found ESs ranging from 0.22 (instructional explanations /

worked examples, Wittwer & Renkl, 2010) to 0.96 (self-regulation interventions in primary mathematics, Dignath & Büttner, 2008).

However, there is “considerable variation” (Higgins et al., 2005, p.34) across approaches and studies. Implementing these approaches is not straightforward. The development of thinking skills, metacognition and self-regulation takes time and the role of the teacher is important in ensuring a careful match between the approach, the learner and the subject (Higgins et al., 2005). Lai’s (2011) review recommends that learners are exposed to a variety of explicitly taught strategies, urging teachers to promote metacognitive processes through modelling or scaffolding a strategy while simultaneously verbalizing their thinking or asking questions of the learners to highlight aspects of the strategy. Teachers need to be careful that the strategy use does not detract from the mathematical task (Rittle-Johnson et al., 2017).

Regardless of the strategy being taught explicitly, learners need significant time to imitate, internalise and independently apply strategies, and they need to experience the same strategies being used repeatedly across many lessons (Ellis et al., 2014). These findings are supported by two meta-analyses. Dignath & Büttner (2008) found that, at primary school level, ESs increased with the number of training sessions, while Xin & Jitendra (1999) found that long-term interventions produced substantially higher ESs ($d=2.51$ for long-term interventions compared with $d=0.73$ for intermediate-length interventions). It is likely that the time required is significantly greater than for other concepts, without the ‘drop-off’ seen with approaches such as the use of manipulatives.

One thinking skills programme developed in England, Cognitive Acceleration in Mathematics Education (CAME), appears to be particularly promising. Higgins et al.’s (2005) synthesis focused on the effects of thinking skills programmes and found that thinking skills approaches may have a greater effect on attainment in mathematics (and science) than they do on reading (ES for mathematics $d=0.89$ compared to English $d=0.48$), although the difference was not significant. Higgins et al. included four studies of the effects of Cognitive Acceleration, a programme that has been extensively used in England, and found an immediate effect on attainment of $d=0.61$. However, these studies were set either in science or in early-years education. Several studies of the CAME programme show very promising results. One quasi-experimental study of the CAME programme delivered in Years 7 and 8 found a relatively large effect of $d=0.44$ on GCSE grades in mathematics three years after the end of the intervention (Shayer & Adhami, 2007). Another study of the CAME programme delivered in Years 1 and 2 found a medium effect of $d=0.22$ on Key Stage 2 mathematics. Finally, a study of the programme delivered in Years 5 and 6 found an immediate effect on Key Stage 2 mathematics of $d=0.26$.

Strategies that encourage elaboration and self-explanation appear to be beneficial. Elaboration involves students explaining mathematics to someone else, often in a collaborative learning situation, drawing out connections with previous learning (Kramarski & Mevarech, 2003). Self-explanation involves learners elaborating for themselves, rather than for a public audience. Both have links to the use of worked examples (providing a detailed example of a solution to a task/problem which is then used on similar tasks/problems). Wittwer & Renkl (2010) found that in mathematics, worked examples, *in combination with instructional guidance*, appeared to be effective, with the ES for mathematics $d = 0.22$ (95%CI 0.06, 0.38), and to be effective in developing conceptual understanding. However, providing instructional guidance appears to be no better than encouraging self-explanation.

There is some evidence to suggest that, in primary, focusing on cognitive strategies may be more effective, whereas, in secondary, focusing on learner motivation may be important. Dignath et al.'s (2008) meta-analysis aimed to better understand the variation in effects and to investigate the impact of various characteristics of different approaches and teaching methods. They found greater effects for self-regulation interventions in mathematics at primary (higher than reading/writing), $d=0.96$ (95%CI 0.71, 1.21) than at secondary, $d=0.23$ (95%CI 0.07, 0.38), which was lower than reading. Coding interventions as cognitive, metacognitive, or motivational, they found that cognitive approaches had the strongest effects in primary mathematics, whereas for secondary mathematics, motivational approaches had a greater effect.³ At secondary level, effects were also stronger where group work was used as a teaching approach.

In response to the variation noted earlier, we found repeated calls across the syntheses for more robust studies, for clear definitions of terms, and for stronger outcome measures relying less on self-reported scales (Gascoine et al., 2017). Furthermore, Higgins et al. (2005) note the need for improved reporting, ensuring that methodological details and results crucial to later systematic syntheses are not omitted at the reporting or publishing stages.

Executive function – particularly working memory – is known to be associated with mathematics attainment. The overall correlation between executive function and mathematical attainment is $r = 0.31$, 95% CI [0.26, 0.37] (Jacob & Parkinson, 2015), while overall correlations between WM and mathematical attainment are reported by Jacob & Parkinson (2015) as $r = 0.31$, 95% CI [0.22, 0.39] and by Peng et al. (2016) as $r = .35$, 95% CI [.32, .37]. Although executive function / working memory appear to be correlated, both with general attainment and with mathematics, there is no evidence of a causal relationship (Jacob & Parkinson, 2015, p. 512; see also Friso-van den Bos et al., 2013). Across mathematical domains, Peng et al. (2016) found the strongest correlations for WM with word-problem solving.

In terms of Working Memory Training (WMT), Jacob & Parkinson (2015) found across five intervention studies no compelling evidence that impacts on executive function lead to increases in academic achievement. In mathematics, Melby-Lervåg & Hulme (2013) found small and non-significant effects of WMT on arithmetic ($d=0.07$), while Schwaighofer et al. (2015), building on Melby-Lervåg & Hulme's analysis, found little evidence of short-term ($d=0.09$) or long-term (0.08) transfer of WMT to mathematical abilities.

Finally, there is a need for collaborations between mathematical cognition, learning scientists and mathematics educators in order to make sense of the growing, and somewhat varied, corpus of research in this area. There is a need to understand whether there is a causal link between executive function and mathematics achievement, prior to interventions designed to improve executive function in school-age children being piloted and scaled-up.

³ Cognitive strategies involve rehearsal, elaboration and organisation skills such as underlining, summarising and ordering (Dignath et al., 2008, p.236). They are essentially the linchpins of the later metacognitive strategies of planning, monitoring and evaluating, and a part of the continuum of children's metacognitive development, which is known to be age-related (Ellis et al., 2012; Gascoine et al., 2017; Lai, 2011).

Evidence base

On the teaching of thinking skills, metacognition and/or self-regulation, we drew on six meta-analyses and one systematic review. These six meta-analyses synthesised a total of 233 studies published between 1981 and 2015 and predominantly are judged to be of high methodological quality. There was little or no overlap in the studies included when judged against the largest meta-analysis (Dignath & Büttner, 2008).

On working memory training, we drew on two meta-analyses, both of high quality, with no overlap in the studies synthesised.

Meta-analysis (Metacognition)	<i>k</i>	Quality	Date Range	% overlap with Dignath & Büttner (2008)
Dignath & Büttner (2008)	49 primary 25 secondary	2	1992-2006	N/A
Donker et al. (2014)	58	3	2000-2012	7% (4/58)
Higgins et al. (2005)	29	3	1984-2002	3% (1/29)
Rittle-Johnson et al. (2017)	26	2	1998-2015	4% (1/26)
Wittwer & Renkl (2010)	21	3	1985-2008	0% (0/21)
Xin & Jitendra (1999)	25	3	1981-1995	0% (0/25)

Meta-analysis (Working memory)	<i>k</i>	Quality	Date Range
Melby-Lervåg & Hulme (2013)	23	3	2002-2011
Schwaighofer et al. (2015)	47	3	2002-2014

Directness

Our overall judgement is that the available evidence is of generally high directness.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	Included studies are worldwide, but many come from the UK or US. For example, in Higgins et al.'s meta-analysis, over half the included studies were from the UK or US.

How the intervention was defined and operationalised	2	Definition of variables and explanation of strategies may be a threat to directness. Multiple models and constructs exist and it is not always clear where the boundaries to a particular construct lie.
Any reasons for possible ES inflation	3	
Any focus on particular topic areas	3	
Age of participants	3	Metacognitive thinking is now accepted as beginning in children as young as three, and the studies included here reflect the full age-range under consideration.

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
Interventions and Training Effects			
<i>Thinking Skills Intervention on Mathematical Attainment</i>			
Higgins et al. (200	0.89 [0.50, 1.29]	k=9	The overall cognitive effect size was 0.62 (k=29). The overall effect size (including cognitive, curricular and affective measures) was 0.74. There was relatively greater impact on tests of mathematics (0.89) and science (0.78), compared with reading (0.4).
<i>Self-Regulation Interventions on Mathematics Attainment</i>			
Dignath & Büttner (2008); <i>Primary mathematics</i>	0.96 [0.71, 1.21]	49 primary 28 mathematics (primary & secondary combined)	Higher effect than for reading (0.44). “Effect sizes for mathematics performance at primary school were higher: for interventions focusing on cognitive strategy instruction (reference category) rather than on metacognitive reflection (B=-1.08) for interventions with a large number of sessions (B=0.05)” (p. 247)
Dignath & Büttner (2008); <i>Secondary mathematics</i>	0.23 [0.07, 0.38]	25 secondary 28 mathematics (primary &	Lower effect than for reading (0.92). “Effect sizes representing mathematics performance at secondary school were higher:

		secondary combined)	<p>if the theoretical background of the intervention focused on motivational ($B=0.55$) rather than on metacognitive learning theories (reference category). No significant difference was found compared to social-cognitive theories.</p> <p>if group work was not used as a teaching method (constant) rather than if it was used ($B=-0.65$).</p> <p>with an increasing number of training sessions ($B=0.02$).” (pp. 247-8)</p>
<i>Self-Explanation Prompts on Mathematical Attainment</i>			
Rittle-Johnson et al. (2017); <i>procedural knowledge</i>	0.28 [0.07 0.49]	19	Immediate post-test. Delayed post-test ES = 0.13 [-0.13 0.39]
Rittle-Johnson et al. (2017); <i>conceptual knowledge</i>	0.33 [0.09 0.57]	16	Immediate post-test. Delayed post-test ES = -0.05 [-0.29 0.19]
Rittle-Johnson et al. (2017); <i>procedural transfer</i>	0.46 [0.16 0.76]	9	Immediate post-test. Delayed post-test ES = 0.32 [0.02 0.63]
<i>Metacognition and self-regulation on Mathematical Attainment</i>			
Donker et al. (2014)	0.66	58 Studies, 44 interventions in mathematics	<p>Overall attainment = .66 (SE = .05, 95%CI .56 to .76)</p> <p>Writing = 1.25</p> <p>Reading = 0.36</p> <p>These domains differed in terms of which strategies were the most effective in improving academic performance. However, metacognitive knowledge instruction appeared to be valuable in all of them.</p>
<i>Instructional Explanations (Worked Examples) on Mathematics Attainment</i>			
Wittwer & Renkl (2010)	0.22 [0.06, 0.38]	14	The weighted mean (across subjects) effect size of 0.16 [0.03, 0.30] was small but statistically significant, $p=0.04$. Two other subjects were examined (science and instructional design) – mathematics was significantly different from instructional design

			but not from science. Science: 0.21 [-0.02, 0.44] Instructional design: -0.28 [-0.71, 0.16]
<i>Strategy training (incorporating explicit instruction and/or metacognitive strategies) in word-problems in mathematics for students with learning disabilities</i>			
Xin & Jitendra (1999)	0.74 95% CI [0.56, 0.93]	12	This compares with other forms of instruction: Representation (k=6) d=1.77, 95%CI [1.43, 2.12] CAI (k=4) d=1.80 95%CI [1.27, 2.33] Other (k=5) d=0.00 95%CI [-0.26, 0.26]
<i>Thinking Skills Intervention on Mathematical Attainment</i>			
Higgins et al. (200	0.89 [0.50, 1.29]	k=9	The overall cognitive effect size was 0.62 (k=29). The overall effect size (including cognitive, curricular and affective measures) was 0.74. There was relatively greater impact on tests of mathematics (0.89) and science (0.78), than with reading (0.4).
<i>Working Memory Training on Arithmetic</i>			
Melby-Lervåg & H (2013)	0.07 95% CI [-0.13, 0.27]	7	The mean effect size was small and nonsignificant. All long-term effects of working memory training on transfer measures were small and nonsignificant.
<i>Transfer effect of WM training to mathematical abilities</i>			
Schwaighofer et al (2015); <i>short-term</i>	0.09 [-0.09, 0.27]	15	This analysis builds on Melby-Lervåg & Hulme (2013), examining the near and far transfer of WMT.
Schwaighofer et al (2015); <i>long-term</i>	0.08 [-0.12, 0.28]	8	

References

Meta-analyses included

- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231-264.
- Donker, A. S., De Boer, H., Kostons, D., van Ewijk, C. D., & Van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1-26. DOI: 10.1016/j.edurev.2013.11.002
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Melby-Lervåg, M. m., & Hulme, C. c. (2013). Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology*, 49(2), 270-291.
- Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM*, 1-13.
- Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does Working Memory Training Transfer? A Meta-Analysis Including Training Conditions as Moderators. *Educational Psychologist*, 50(2), 138-166. doi: 10.1080/00461520.2015.1036274
- Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review*, 22(4), 393-409.
- 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(4), 207-225.

Meta-analyses excluded

- Durkin, K. (2011). *The self-explanation effect when learning mathematics: A meta-analysis*. Presented at the Society for Research on Educational Effectiveness 2011, Available online [accessed 13th March 2017]: <http://eric.ed.gov/?id=ED518041>. [Full paper not available – appears to be same study as Rittle-Johnson et al. (2017)]

Secondary Meta-analyses

- Friso-van den Bos, I., van der Ven, S. H., Kroesbergen, E. H., & van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational research review*, 10, 29-44.
- Jacob, R., & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement. *Review of Educational Research*, 85(4), 512-552. doi:10.3102/0034654314561338
- Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory

domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, 108(4), 455-473. doi:10.1037/edu0000079

Systematic Reviews

Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, 5(1), 3-57.

Other references

Adhami, M. (2002). Cognitive acceleration in mathematics education in years 5 and 6: Problems and challenges. In M. Shayer & P. S. Adey (Eds.), *Learning Intelligence: Cognitive Acceleration across the Curriculum from 5 to 15 years* (pp. 98-117). Buckingham: Open University Press.

Ellis, A. K., Bond, J. B., & Denton, D. W. (2012). An analytical literature review of the effects of metacognitive teaching strategies in primary and secondary student populations. *Asia Pacific Journal of Educational Development (APJED)*, 1(1), 9-23.

Ellis, A. K., Denton, D. W., & Bond, J. B. (2014). An analysis of research on metacognitive teaching strategies. *Procedia-Social and Behavioral Sciences*, 116, 4015-4024.

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal*, 40(1), 281-310.

Lai, E. R. (2011). *Metacognition: A literature review*. Always learning: Pearson research report.

Shayer, M., & Adhami, M. (2007). Fostering Cognitive Development Through the Context of Mathematics: Results of the CAME Project. *Educational Studies in Mathematics*, 64(3), 265-291.

Shayer, M., & Adhami, M. (2010). Realizing the cognitive potential of children 5–7 with a mathematics focus: Post-test and long-term effects of a 2-year intervention. *British Journal of Educational Psychology*, 80(3), 363-379. doi:10.1348/000709909x482363

7 Resources and Tools

7.1 Calculators

What are the effects of using calculators to teach mathematics?

Calculator use does not in general hinder students' skills in arithmetic. When calculators are used as an integral part of testing and teaching, their use appears to have a positive effect on students' calculation skills. Calculator use has a small positive impact on problem solving. The evidence suggests that primary students should not use calculators every day, but secondary students should have more frequent unrestricted access to calculators. As with any strategy, it matters how teachers and students use calculators. When integrated into the teaching of mental and other calculation approaches, calculators can be very effective for developing non-calculator computation skills; students become better at arithmetic in general and are likely to self-regulate their use of calculators, consequently making less (but better) use of them.

Strength of evidence: HIGH

Findings

Two meta-analyses, Ellington (2003) and Hembree & Dessart (1986), synthesised studies of handheld calculator use. Both meta-analyses found that calculator use did not hinder students' development of calculation skills when tested without calculators, and may have had a small positive effect in some areas of mathematics. However, when calculators were permitted in the testing as well as the teaching, calculator use was found to have a positive effect on students' calculation skills. In addition, both meta-analyses found small positive effects of calculator use on students' problem solving. Ellington suggests that the increase in problem-solving skills "may be most pronounced ... when special curriculum materials have been designed to integrate the calculator in the mathematics classroom" (p. 456). Both meta-analyses found that students taught with calculators had more positive attitudes to mathematics.

A large-scale research and development project in England, the Calculator-Aware Number (CAN) project provides further evidence in the English context (Shuard et al., 1991). In a follow-up study examining the effects of a "calculator aware" curriculum on students who had experienced calculators throughout their primary schooling, Ruthven (1998) found that, compared to a control group, students' understandings of and fluency with arithmetic were greater. A key paragraph in Ruthven (1998) states:

In the post-project schools, pupils had been encouraged to develop and refine informal methods of mental calculation from an early age; they had been explicitly taught mental methods based on 'smashing up' or 'breaking down' numbers; and they had been expected to behave responsibly in regulating their use of calculators to complement these mental methods. In the non-project schools, daily experience of 'quickfire calculation' had offered pupils a model of mental calculation as something to be done quickly or abandoned; explicit teaching of calculation had emphasised approved written methods; and pupils had little experience of regulating their own use of calculators. (pp. 39-40)

In addition, the intervention group students used calculators less often, and mental methods more often, than the control group.

Hembree and Dessart (1986) found that, at Grade 4 (Year 5 in England), in contrast to other grades, calculator use had a negative effect. This strikes a cautionary note, and Hembree and Dessart comment that “calculators, though generally beneficial, may not be appropriate for use at all times, in all places, and for all subject matters” (p. 25). In an analysis of TIMSS 2007 data, Hodgen (2012) found that, at Year 5, the attainment of students in countries where calculator use was unrestricted was significantly lower than it was in those countries where calculator use was either restricted or banned. However, the reverse was true for Year 9: the attainment of students where calculator use was unrestricted was higher than it was for those where it was banned. The Leverhulme Numeracy Research Programme also identified different effects from different types of calculator use. Brown et al. (2008) found that allowing students access to calculators either rarely or on most days was negatively associated with attainment. This suggests that calculators should be used moderately but not excessively, and for clear purposes, particularly at primary. As found in the CAN project (Shuard et al., 1991), calculators need to be used proactively to teach students about number and arithmetic alongside the teaching of mental and pencil-and-paper methods; students also benefit from learning to make considered decisions about when, where and why to use different methods. Indeed, in a retrospective analysis of cumulative evidence about CAN, Ruthven (2009) argued that how calculators are used and integrated into teaching is crucial. This analysis supports a principled approach to the use of calculators, in which students are taught, for example, estimation and prediction strategies that they can use to check and interpret a calculator display.

The meta-analyses did not distinguish between basic and scientific calculators. However, 22 of Ellington’s 54 studies (41%) focused on graphic calculators, and moderator analysis found that graphic calculators had higher effects for testing with calculators, problem-solving and attitudes to maths, although there was considerable variation in these effects.

Evidence base

We identified two meta-analyses synthesising a total of 133 studies over the period 1969-2002: Ellington (2003): 54 studies (methodological quality: high), and Hembree & Dessart (1986): 79 studies (methodological quality: medium). Ellington (2003) builds explicitly on Hembree & Dessart and takes a very similar theoretical frame. The results of the two meta-analyses are consistent, although more weight should be placed on Ellington’s more recent study, because there have been significant changes in the use, availability, functionality and student familiarity of calculators since Hembree & Dessart’s search period (1969-1983). A third meta-analysis (Smith, 1986) was excluded due to extensive overlap with the studies included in Ellington’s meta-analysis.

The majority of included studies in Ellington’s (2003) meta-analysis examined students’ acquisition of skills as measured by immediate post-tests. Too few studies examined retention (through delayed post-testing) or transfer (to calculator use in other subject domains) for conclusions to be drawn, and further research is needed in these areas. Whilst the findings of the two meta-analyses are consistent, Ellington’s moderator analysis indicates a relatively high degree of unexplained variation.

Directness

The majority of the studies included in both Ellington's (2003) and Hembree & Dessart's (1986) meta-analyses were conducted in the US. Nevertheless, these findings are judged to apply to the English context, which is supported by the evidence from the CAN project (Shuard et al., 1991). Further, CAN suggests some general principles that can be applied to the classroom use of calculators, although, as Shuard et al. observed, "a calculator-aware number curriculum is much more than a conventional number curriculum with calculator use 'bolted on'. Nor is it a wholly 'calculator-based' one. ... such an approach requires careful planning, particularly of curriculum sequences to underpin continuity and progression in children's learning." (p. 13).

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	The studies in both meta-analyses were conducted in the US. In the absence of reasons to the contrary, these findings are judged to apply to England. A large-scale study at primary provides further weight to this.
How the intervention was defined and operationalised	3	The meta-analyses focus on calculator use as a general strategy rather than on any particular interventions. The research suggests some general and applicable principles for the use of calculators.
Any reasons for possible ES inflation	3	
Any focus on particular topic areas	3	The focus is on calculation and problem-solving, which are central to the research question.
Age of participants	3	The meta-analyses cover the 8-13 age range (and beyond).

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Quality	Comments
<i>Effect of calculator use on calculation skills in tests where calculators were not permitted</i>				
Ellington (2003)	-.02	14	3	Computational aspects of operational skills reported. Ellington found various additional ESs, which vary between $g = -.05$ (conceptual aspects) and $g = .17$

				(operational skills overall).
Hembree & Dessart (1986)	.137	57	2	ES for Grade 4: $g = -.152$ ($k=7$, $p<.05$). Low attainers ES $g = -.107$ ($k=13$, n.s.)
<i>Effect of calculator use on calculation skills in tests where calculators were permitted</i>				
Ellington (2003)	.32	19	3	Operational skills reported. Ellington found various additional ESs, which vary between $g = .41$ (computational aspects) and $g = .44$ (conceptual aspects).
Hembree & Dessart (1986)	.636	29	2	Computational aspects reported; overall operational skills $g = .737$, but studies found to be heterogeneous.
<i>Effect of calculator use on problem-solving</i>				
Ellington (2003)	.22	12	3	
Hembree & Dessart (1986)	.203	33	2	Operational skills overall for "other" grades (i.e., not G4 or G7) of effect on problem-solving without calculators. Various other ESs found that vary between $g = .005$ and $g = .458$. ESs for problem solving with calculators higher.
<i>Effect of calculator use on attitudes to mathematics</i>				
Ellington (2003)	.20	12	3	
Hembree & Dessart (1986)	.190	56	2	

References

Meta-analyses included

Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, 34, 433-463.

Hembree, R., & Dessart, D. J. (1986). Effects of hand-held calculators in precollege mathematics education: A meta-analysis. *Journal for research in mathematics education*, 17(2), 83-99.

Meta-analyses excluded

Ellington, A. J. (2006). The effects of non-CAS graphing calculators on student achievement and attitude levels in mathematics: A meta-analysis. *School Science and Mathematics*, 106(1), 16-26. [Excluded because very few of the studies covered the 8-12 age range.]

Smith, B.A. (1996). A meta-analysis of outcomes from the use of calculators in mathematics education. (Doctoral dissertation, Texas A & M University-Commerce, 1996). Dissertation Abstracts International, 58(03), 787.
[Excluded due to overlap with Ellington, 2003: 87.5%]

Other references

Brown, M., Askew, M., Hodgen, J., Rhodes, V., Millett, A., Denvir, H., & Wiliam, D. (2008). Individual and cohort progression in learning numeracy ages 5-11: Results from the Leverhulme 5-year longitudinal study. In A. Dowker (Ed.), *Children's Mathematical Difficulties: Psychology, Neuroscience and Education* (pp. 85-108). Oxford: Elsevier.

Hodgen, J. (2012). Computers good, calculators bad. In P. Adey & J. Dilllon (Eds.), *Bad Education: Debunking educational myths*. Maidenhead: Open University Press.

Ruthven, K. (2009). Towards a calculator-aware mathematics curriculum. *Mediterranean Journal for Research in Mathematics Education*, 8(1), 111-124.

Ruthven, K. (1998). The Use of Mental, Written and Calculator Strategies of Numerical Computation by Upper Primary Pupils within a 'Calculator-Aware' Number Curriculum. *British Educational Research Journal*, 24(1), 21-42.

Shuard, H., Walsh, A., Goodwin, J., & Worcester, V. (1991). *Calculators, children and mathematics: The Calculator-Aware Number curriculum*. Hemel Hempstead: Simon & Schuster.

7.2 Technology: technological tools and computer-assisted instruction

What is the evidence regarding the use of technology in the teaching and learning of maths?

Technology provides powerful tools for representing and teaching mathematical ideas. However, as with tasks and textbooks, how teachers use technology with learners is critical. There is an extensive research base examining the use of computer-assisted instruction (CAI), indicating that CAI does not have a negative effect on learning. However, the research is almost exclusively focused on systems designed for use in the US in the past, some of which are now obsolete. More research is needed to evaluate the use of CAI in the English context.

Strength of evidence (Tools): LOW

Strength of evidence (CAI): MEDIUM

Findings

We identified 11 meta-analyses addressing aspects of technology. Despite this relatively large evidence base, we judge the evidence regarding technology to be limited. The 11 meta-analyses were published between 1977 and 2017 and synthesise studies published between 1967 and 2016. During this period, there have been very dramatic changes in the scope, capability, availability and familiarity of technology. The term ‘technology’ has expanded to cover a wide range of very different applications and devices, each of which may have different potential uses in the teaching and learning of mathematics. Several of the meta-analyses aggregated the effects of different uses of technology, indicating ESs of $d=0.28$ in general (Li & Ma, 2010) and $d=0.47$ for primary (Chauhan, 2017). However, the diverse range of technologies synthesised in each of these meta-analyses makes interpretation of the effects problematic, beyond a general effect for innovation and novelty. In order to address this diversity, we present our findings under two categories:

Technological tools: A vast range of technological hardware and software is used in mathematics classrooms in England. This is sometimes referred to as digital technology or ICT (information and communication technology), and in this module we refer to these as technological tools. The tools addressed in the meta-analyses are a subset of these, and include mobile devices, dynamic geometry software, exploratory computer environments and educational games.

Computer-assisted instruction (CAI): CAI covers a broad range of computer-based systems designed to deliver all or part of the curriculum or to support the management of learning by providing assessment and feedback to learners. Some CAI is designed to supplement regular teaching, whilst other CAI is comprehensive. CAI is intended to be adaptive to the needs of individual learners, and one meta-analysis focuses on Intelligent Tutoring Systems [ITS], which have ‘enhanced adaptability’ and attempt to replicate human tutoring.

Note: Calculators are considered in a separate module, because the evidence base is substantial and has a specific focus on calculation and arithmetic.

Technological Tools

Four meta-analyses examined the effects of using technological tools on attainment in comparison to non-use, and one meta-analysis looked at the effect on learner attitudes. A very large ES was reported for dynamic geometry software ($d=1.02$)

(Chan & Leung, 2014), but this is likely to have been inflated by the exploratory nature of the study. The ESs reported for other exploratory approaches, game-based approaches and hand-held devices were medium to small: exploratory computer environments $d=0.60$ (Sokolowski et al., 2015), game-based approaches $d=0.26$ (Tokac et al., 2015) and the use of mobile devices, $d=0.16$ (Tingir et al., in press). Technological tools have the potential for large effects, but, whilst Chan and Leung's finding suggests that the use of DGS has considerable potential, more substantial research is needed before assuming that dynamic geometry software will be transformative in the classroom.

One meta-analysis of the impact of the use of technological tools on learners' attitudes towards mathematics reported an ES of 0.35 (Savelsburgh et al., 2016). Moderator analysis also revealed that the impact on attitude lessened as learners got older, although this may be affected by a general tendency for attitudes to become more negative with age through the school years.

Li and Ma (2010) stress that how technology is used matters. Two best-evidence studies by Slavin et al. (2008, 2009) indicate that technology applications appear to produce lower effects than interventions aimed at changing teaching. As technology advances, there will be an increased need for professional development for teachers to keep pace with this change (Chauhan, 2017).

Computer-Assisted Instruction

There is an extensive research base considering the impact of the use of CAI on mathematics attainment, although it is limited by being largely conducted with systems that were designed some time ago for use in the US. The meta-analyses produce ESs ranging from 0.01 to 0.41. Smaller ESs are reported in the most recent studies, which are of higher methodological quality. Cheung & Slavin (2013) report an ES of 0.16 for CAI, although the effect reduced to a non-significant 0.06 when including only large randomised controlled studies. Steenbergen-Hu & Cooper (2013) found an ES of 0.01 for ITS approaches. Overall the evidence base indicates that the use of CAI does not have a large negative effect on learning and may be a valuable supplement to teaching, which can free the teacher to focus on other aspects of teaching. This supplemental use is supported by findings (Schmid et al., 2009) that the effects of the use of technology are stronger when the technology use is low (ES=0.33) or medium (ES=0.29) compared with high usage (ES=0.14). Ruthven (2001) cites one extensive study conducted in the UK in the 1990s on the effects of integrated learning systems (ILS), which concluded that ILS have shown effectiveness for the development of basic skills, but not for reasoning with numbers. It is likely that the capabilities of CAI, ITS and other ILS systems will develop considerably alongside advances in technology and big data. There is a need for further studies in England to evaluate these developments and to establish which aspects of CAI have the potential to improve learning.

Evidence base

Overall the evidence base is fairly strong, but caution must be applied in an area subject to such rapid change. We have drawn on 11 meta-analyses synthesizing 434 studies covering the period 1962-2016. Study overlap would appear to lie within the usual range; for example, 23% of the studies in Steenbergen-Hu & Cooper (2013) overlap with the 64 studies included in Cheung & Slavin (2013).

The meta-analyses represent a range of methodological quality. In particular, many of the primary studies reviewed in the meta-analyses of technology tools are exploratory studies and many are small-scale and without pre-tests. Conversely, there is a very extensive programme of research on CAI with large-scale RCTs, but these are US-based, and research is needed to understand how they might work in the English context.

Moderator analyses included within nine of the meta-analyses suggest that elementary and/or middle school learners return similar or higher ESs than do secondary-age learners. Cheung & Slavin (2013) note this to be consistent with previous reviews. Over half of the included meta-analyses looked at the time-span of the intervention. As Table 1 shows, the results indicate a range of ESs, with no clear picture as to the 'best' intervention length.

Directness

Technology tools

As Li & Ma (2010) observe, context matters in the use of technology tools: "The effectiveness of mathematics learning with technology is highly dependent on many other characteristics such as teaching approaches, type of programs, and type of learners." (p. 200) Whilst this is the case for any broad set of tools (e.g., manipulatives), the technology area is particularly broad. The range and uses of technology tools has changed, and continues to change, rapidly. Hence, many of the tools examined are innovative and novel. Novelty may affect implementation positively, because the novelty may motivate learners and teachers. Novelty can also affect implementation negatively, because teachers may have difficulty using technology through lack of expertise or guidance.

Technology applications

The stronger primary studies of CAI are largely conducted in the US and evaluate dated CAI systems. None of the CAI studies were in England or with programmes designed for the English context (although some technology applications designed for the English context do exist).

The recent UK trial of *PowerTeaching Maths* (Slavin et al, 2013) demonstrates that the transfer of a US technology-focused intervention to the context of English classrooms is not straightforward. *PowerTeaching Maths* is a technology-enhanced teaching approach based on cooperative learning in small groups. The researchers found that implementation was limited by the prevalence of within-class ability grouping in England.

Cheung & Slavin's (2013) findings about large-scale RCTs suggest that, when implemented at scale, the effects of technology applications are likely to be small or negligible, but not negative.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	The greatest threat to directness is the publication date of the included studies, given the speed of technological change.
How the intervention was defined and operationalised	3	Technology tools are generally well-defined, although the ways in which these tools are used is less so.

		Technology applications are largely designed for use in the US curriculum.
Any reasons for possible ES inflation	2	Possible novelty factor.
Any focus on particular topic areas	3	N/A
Age of participants	3	Majority of meta-analyses covered the K-12 range, two covered elementary or elementary and middle school grades. Moderator analysis allowed for exploration of grade-level implications.

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Quality	Comment
<i>Effect of technology use in general on mathematical attainment</i> [NOTE: These meta-analyses combine technology tools and CAI.]				
Li & Ma (2010): Mathematics [1991-2005]	0.28 [0.13, 0.43]	46	2	Studies contained 85 ESs. Interventions with durations of more than 1 year had lower effects than those of one term.
Chauhan (2017): Primary (elementary) [2000-2016]	0.47 [0.35, 0.59]	41	2	ES for mathematics reported. Overall $d=0.55$ across subjects, $k=122$. This is a general meta-analysis and, hence, the data extraction for mathematics-specific instructional features is limited. ES may be inflated because more than half of the included studies have no pre-test.
<i>Effect of Computer Aided Instruction (CAI) and Intelligent Tutoring Systems (ITS) on mathematical attainment</i>				
Cheung & Slavin (2013): CAI (including ITS) [1980-2010]	0.16 [0.11, 0.20]	74	3	Computer-Assisted Instruction (CAI), and includes Intelligent Tutoring Systems (ITS). Applications were categorised as supplemental, computer-managed (or assessment-based systems) or comprehensive. Supplemental was found to have larger effects (and to have a more extensive evidence base). The focus of this meta-analysis is on “replicable programs used in realistic settings over periods of at least 12 weeks” using standardised tests (p. 95). The

				programmes used are all developed for the US. Large randomised controlled studies had smaller (and non-significant) effects $d = 0.06$.
Steenbergen-Hu & Cooper (2013): ITS [1997-2010]	0.01 [-0.10, 0.12]	26 reports, 31 studies comparing ITS to regular classroom instruction, 17 studies with adjusted effects	3	Some studies had no pre-test. ITS had a negative effect on low-attaining learners (only significant on a fixed effects model, $g = -0.19$, $k=3$). This result needs to be treated with caution since it is based on a small number of studies and the effect is only significant on some models.
Kuchler (1998): CAI [1976-1996]	0.28	61	2	Weighted ES with 4 outliers removed. No CI given for ES
Hartley (1977): CAI [1967-1976] Synthesised studies from 1962, but first study involving technology (CAI) dated 1967.	0.41 [0.29, 0.53]	89	1	CI calculated from standard error (by review authors). At Grade 5 (Y6), ES were larger for low attainers compared to high attainers, and also larger for 1 session per week compared to daily (5) sessions per week (p.81). Effects appear to decrease with age, although younger (Y3) and older (Y12) groups are out of our age range.
<i>Effect of technology tools on mathematical attainment</i>				
Sokolowski et al. (2015): Exploratory Computer Environments [2000-2013]	0.60 [0.53, 0.66]	24 primary ESs	2	Meta-analysis includes a very broad range of packages under the umbrella of Exploratory Computer Environments (DGS, games, generic gaming, collaborative software) and a broad range of different approaches.
Chan & Leung (2014):	1.02 [0.56, 1.48]	9	2	Short-term instruction with DGS significantly improved the achievement of primary learners $d =$

Dynamic Geometry Software [2002-2012]				1.82 [1.38, 2.26], $k=3$. The effect size may be inflated, because studies were largely small scale and of short duration.
Tingir et al. (In press): Mobile devices [2010-2014]	0.16 [- 0.55, 0.87]	3	3	The effect for mathematics is not significant, but this is based on a very small sample of studies ($k=3$). Effects in science and reading were larger (and reading was significant).
Tokac et al. (2015): Game based learning [2000-2011]	0.26 [0.01, 0.50]	13	2	This conference paper reports work in progress, although some detail is provided. Moderator analysis was carried out but is not reported in detail.
<i>Effect of use of technological tools on attitude</i>				
Savelsburgh et al (2016) 'innovative' ICT-rich environments [1988-2014]	0.35 [0.24, 0.47]	11	3	Effect of innovative mathematics and science teaching on attitudes. Innovative approaches include ICT-rich environments (19 of 65). Most of the ICT-rich studies were conducted in mathematics education (11 of 19). No difference found for mathematics or for ICT-rich environments, but effects decrease with age.
<i>Effect of technology tools on learning of algebra</i>				
Haas (2005)	0.07	7	2	
Rakes et al. (2010)	0.17	23	3	

References

Meta-analyses included

- Chan, K. K., & Leung, S. W. (2014). Dynamic geometry software improves mathematical achievement: Systematic review and meta-analysis. *Journal of Educational Computing Research*, 51(3), 311-325.
- Chauhan, S. (2017). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education*, 105, 14-30.
- 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.
- Hartley, S. S. (1977) Meta-Analysis of the Effects of Individually Paced Instruction In Mathematics. Doctoral dissertation University of Colorado at Boulder.
- Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin*, 89(642), 24-46.

- Kuchler, J. M. (1998) The effectiveness of using computers to teach secondary school (grades 6-12) mathematics: A meta-analysis. Ph.D. thesis, University of Massachusetts Lowell.
- Li, Q. and Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, 22(3), 215-243.
- Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of Instructional Improvement in Algebra A Systematic Review and Meta-Analysis. *Review of Educational Research*, 80(3), 372-400.
- Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, 19, 158-172.
- Sokolowski, A., Li, Y., & Willson, V. (2015). The effects of using exploratory computerized environments in grades 1 to 8 mathematics: a meta-analysis of research. *International Journal of STEM Education*, 2(1), 1-17.
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970-987.
- Tingir, S., Cavlazoglu, B., Caliskan, O., Koklu, O., Intepe-Tingir, S. (In press) Effects of mobile devices on K-12 students' achievement: A meta-analysis. *Journal of Computer Assisted Learning*, DOI: 10.1111/jcal.12184
- Tokac, U., Novak, E. & Thompson, C. (2015). Effects of Game-Based Learning on Students' Mathematics Achievement: A Meta-Analysis. Representing Florida State University with a poster presentation at 2015 Statewide Graduate Student Research Symposium, University of Central Florida, Orlando, FL, April 24, 2015.

Secondary meta-analyses

- Slavin, R. E., & Lake, C. (2008). Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis. *Review of Educational Research*, 78(3), 427-515. doi:10.3102/0034654308317473. [This focuses on the effectiveness of specific programmes designed for use in the US, some of which are no longer commercially available. The analysis of CAI in general has been superseded by Cheung & Slavin (2013). Used to quantify comparison to instructional programmes: "Median effect sizes for all qualifying studies were +0.10 for mathematics curricula, +0.19 for CAI programs, and +0.33 for instructional process programs." (p. 476)]
- Slavin, R. E., Groff, C., & Lake, C. (2009). Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis. *Review of Educational Research*, 79(2), 839-911. [This focuses on the effectiveness of specific programmes designed for use in the US, some of which are no longer commercially available. The analysis of CAI in general has been superseded by Cheung & Slavin (2013). Used to quantify comparison to instructional programmes: "The weighted mean ES for math curricula was only +0.03. Corresponding numbers were +0.10 for CAI studies and +0.18 for instructional process studies. Among the instructional process programs, however, there

was great variation. Two cooperative learning programs, STAD and IMPROVE, had very positive outcomes (weighted mean ESs of +0.42 and +0.52, respectively), and several other types of approaches had positive effects in one or two studies.” (p. 882)]

Meta-analyses excluded [and reason]

- Chen, H. (2004). The efficacy of mathematics interventions for students with learning disabilities: A meta-analysis. (Order No. 3157959, The University of Iowa). [Excluded due to focus on students with learning difficulties.]
- Demir, S., & Basol, G. (2014). Effectiveness of Computer-Assisted Mathematics Education (CAME) over Academic Achievement: A Meta-Analysis Study. *Educational Sciences: Theory and Practice*, 14(5), 2026-2035. [Too little information contained in paper]
- Dennis, M. S., Sharp, E., Chovanes, J., Thomas, A., Burns, R. M., Custer, B., & Park, J. (2016). A Meta-Analysis of Empirical Research on Teaching Students with Mathematics Learning Difficulties. *Learning Disabilities Research & Practice*, 31(3), 156-168. [Excluded due to focus on students with learning difficulties.]
- Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and special education*, 24(2), 97-114. [Excluded due to focus on students with learning difficulties.]
- Lee, D. S. (2000). A meta-analysis of mathematics interventions reported for 1971-1998 on the mathematics achievement of students identified with learning disabilities and students identified as low achieving. Doctoral Thesis, University of Oregon ProQuest UMI 9963449 [Excluded due to focus on students with learning difficulties.]
- Sahin, B. (2016) Effect of the use of technology in mathematics course on attitude: A meta analysis study. *Turkish Online Journal of Educational Technology*, (November Special Issue), pp. 809-814 [Information too poorly reported in the paper]

Other references

- Ruthven, K. (2001). British research on developing numeracy with technology. In M. Askew & M. Brown (Eds.), *Teaching and Learning Primary Numeracy: Policy, Practice and Effectiveness. A review of British research for the British Educational Research Association in conjunction with the British Society for Research into Learning of Mathematics* (pp. 29-33). Southwell, Notts: British Educational Research Association (BERA).
- Slavin, R. E., Sheard, M., Hanley, P., Elliott, L., & Chambers, B. (2013). Effects of Co-operative Learning and Embedded Multimedia on Mathematics Learning in Key Stage 2: Final Report. York: Institute for Effective Education.

7.3 Concrete manipulatives and other representations

What are the effects of using concrete manipulatives and other representations to teach mathematics?

Concrete manipulatives can be a powerful way of enabling learners to engage with mathematical ideas, provided that teachers ensure that learners understand the links between the manipulatives and the mathematical ideas they represent. Whilst learners need extended periods of time to develop their understanding by using manipulatives, using manipulatives for too long can hinder learners' mathematical development. Teachers need to help learners through discussion and explicit teaching to develop more abstract, diagrammatic representations. Number lines are a particularly valuable representational tool for teaching number, calculation and multiplicative reasoning across the age range. Whilst in general the use of multiple representations appears to have a positive impact on attainment, the evidence base concerning specific approaches to teaching and sequencing representations is limited. Comparison and discussion of different representations can help learners develop conceptual understanding. However, using multiple representations can exert a heavy cognitive load, which may hinder learning. More research is needed to inform teachers' choices about which, and how many, representations to use when.

Strength of evidence (Manipulatives): HIGH

Strength of evidence (Representations): MEDIUM

Findings

The use of concrete manipulatives has been extensively researched and we identified five meta-analyses. The aggregated ESs present a relatively consistent small to moderate effect, $d=0.39$ (Carbonneau et al., 2013), $d=0.22$ (Holmes, 2013), $d=0.39$ (Domino, 2010) and $d=0.29$ (Sowell, 1989). However, within the earlier meta-analyses (Domino, 2010; Sowell, 1989), there was a very considerable degree of unexplained variation, which may be due to methodological or implementation factors, and one meta-analysis (LeNoir, 1989) found too much variation to report an overall effect.

The most recent meta-analysis, Carbonneau et al.'s (2013), was designed to make sense of this variation. Carbonneau et al. re-examined many of the studies included in previous meta-analyses, focusing specifically on those in which learners were taught how to use the concrete manipulatives, and in which conditions involving concrete manipulatives were compared to teaching involving exclusively abstract mathematical symbols. They examined the effects on retention, problem solving and transfer, as well as on attainment overall.⁴ The effects were higher for retention ($d=0.59$, $k=53$) and problem solving ($d=0.48$, $k=9$) than for transfer ($d=0.13$, $k=13$), although there were many more studies of retention. They found that high levels of instruction were associated with higher effects on overall, retention and problem-solving outcomes, but that the opposite was true for transfer outcomes; here, studies with lower levels of instructional guidance had higher effects. Hence, Carbonneau et

⁴ Retention was defined as an "outcome that required students to solve basic facts" (Carbonneau et al., 2013, p. 388), rather than a delayed post-test measure. Problem-solving was defined as tasks which "students were not explicitly instructed on how to complete" (p.?) and transfer as extending knowledge to a new situation or topic. Justification was also examined, but only two studies addressed this outcome.

al. argue that in general explicit teaching helps learners to establish connections between the concrete manipulatives and the intended mathematical ideas, which in turn facilitates comprehension and understanding. However, if the pedagogical objectives are for learners to transfer knowledge to other areas of mathematics, it may be important to reduce the extent of scaffolding on the use of the manipulatives. However, they caution that more research is needed in this area.

Domino (2010) found no significant differences for learners at different attainment levels. However, Carbonneau et al. (2013) found an age effect: concrete manipulatives had a greater effect for learners aged 3-7 ($d=0.33$) and 7-11 ($d=0.45$) than for older learners ($d=0.16$), which they attribute to their developmental stage (Piaget's concrete operational stage). However, the majority of studies were with the 7-11 age group (38 of 55 studies).

Whilst the earlier meta-analyses (Domino, 2010; LeNoir, 1989; Sowell, 1989) found benefits in long-term use of manipulatives, the results also showed variation. Carbonneau et al.'s (2013) study carefully examined the effect of time, and found that, in general, interventions using manipulatives for up to 45 days had a greater effect than interventions over longer periods. However, Carbonneau et al. caution that more research is needed to better understand the effect of instructional time.

In contrast to concrete manipulatives, we found less, and weaker, evidence about the use of representations. Two of the meta-analyses (Holmes, 2013; Sowell, 1989) examined the effects of virtual or pictorial representations compared to manipulatives and in comparison to abstract teaching, and found no significant differences.

There is a great deal of evidence regarding the importance of representations in the learning of mathematics (see, e.g., Nunes et al., 2009; see also Swan, 2005). Indeed, Nunes et al. (2008) observe that representations are fundamental to mathematics: "Conventional number symbols, algebraic syntax, coordinate geometry, and graphing methods, all afford manipulations which might otherwise be impossible." (p. 9) Consequently, learners need to learn to interpret, coordinate and use different mathematical representations to focus on the relevant relations in specific problems. Ainsworth (2006) argues that the question is not whether multiple representations are effective but rather how and under what circumstances they are more or less effective, and presents a research-based framework outlining ways in which two or more representations can interact during teaching and learning: two representations may *complement* each other by providing different information, or one representation may *constrain* the interpretation of the other, or the combination of two representations may enable learners to construct a deeper conceptual understanding. She notes that multiple representations can exert a heavy cognitive load on learners and argues that, all else being equal, the number of representations presented to learners should be the minimum necessary to achieve the pedagogic objectives. More research is needed on how representations should be used and sequenced.

Finally, we note the particular value of using manipulatives and representations in principled ways for specific topics, such as the importance of the number line in extending learners' understanding of whole numbers to fractions, decimals and percentages (e.g., Siegler et al., 2010). For more details, see Mathematical Topics modules.

Evidence base

We reviewed five meta-analyses, which all focused on concrete manipulatives rather than representations more broadly. These five meta-analyses synthesised more than 150 studies published between 1955 and 2012. Two of the meta-analyses are judged to be of high methodological quality, whilst the other three are judged to be of medium quality. There was a relatively small degree of overlap in the studies included when judged against the most recent and methodologically strongest meta-analysis (Carbonneau et al., 2013). More evidence is required about the level, and type, of instructional guidance that should be provided, particularly relating to problem solving and transfer.

The evidence base on the efficacy of representations is much weaker than for manipulatives. There is a need for a robust meta-analysis examining representations.

There is currently a great deal of interest in England concerning Concrete-Pictorial-Abstract (CPA) approaches to teaching mathematics.⁵ However, we found limited evidence about this approach and identified only one potentially relevant meta-analysis (Hughes et al., 2014). However, we have excluded it from the review. This meta-analysis was concerned with students with learning difficulties and synthesised just two studies addressing the effect of CPA, both of which were conducted by the same team of researchers.

Meta-analysis	#	Focus	k	Quality	Date Range	Overlap with Carbonneau et al. (2013)
Carbonneau et al. (2013)	3	Concrete manipulatives	55	3	1955-2010	N/A
Holmes (2013)	5	Concrete and virtual manipulatives	26	3	1989-2012	19%
Domino (2010)	12	Physical manipulatives at primary	31	2	1991-2009	16%
LeNoir (1989)	29	Manipulatives	45	2	1958-1985	20%
Sowell (1989) ⁶	30	Manipulative materials (includes pictorial)	60	2	Pre-1989	≤60%

Directness

A recent research study in England has resulted in a professional publication focused on the use of manipulatives for the teaching of arithmetic (Griffiths et al., 2016).

⁵ <https://www.ncetm.org.uk/resources/48533>

⁶ Sowell (1989) does not provide a list of the original studies included in her meta-analysis. However, the maximum overlap is calculated using the number of studies published pre-1989 in Carbonneau et al. (2013).

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	Most studies were conducted in the US, but this is not judged to be a threat to directness in this area.
How the intervention was defined and operationalised	2	The studies combine a range of different manipulatives and representations. More research is needed on what the level of support and explicit instruction should be for different learning outcomes.
Any reasons for possible ES inflation	2	Carbonneau et al.'s ES may be inflated by the inclusion of studies using a within-subjects design (23.2%). No statistically significant difference was observed between experimental and quasi-experimental studies.
Any focus on particular topic areas	3	
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Effect of concrete manipulatives on attainment in mathematics</i>			
Carbonneau et al. (2103)	0.39, 95% CI [0.33, 0.44]	55	<p>Inclusion criteria: "Stud[ies] ... compare[d] an instructional technique that used manipulatives with a comparison group that taught math with only abstract math symbols [with] ... no iconic representations ... present. The examined instructional treatments must have provided some form of instruction during which students were able to learn from the manipulatives. ... [S]tudies that required students to work with rulers, scales, or calculators were not included, as these were seen as tools rather than manipulatives." (p. 383).</p> <p>Effect was higher for retention ($d=0.59$, $k=53$) and problem solving ($d=0.48$, $k=9$) than for transfer ($d=0.13$, $k=13$), although there were many more ESs for retention.</p> <p><i>Level of instructional guidance:</i> Overall ($d=0.46$, high, $d=0.29$, low), Retention ($d=0.90$, high, $d=0.19$, low), Problem</p>

			<p>solving ($d=1.06$, high, $d=0.04$, low), Transfer ($d=0.00$, high, $d=0.27$, low).</p> <p><i>Instructional time:</i> ($d=0.34$, ≤ 14 days, $d=0.45$, 15-45 days, $d=0.14$, ≥ 46 days). However, Carbonneau et al cautioned that they were not able to disentangle the instructional time from the study length.</p> <p><i>Age /developmental stage of learners:</i> Age 3-7, pre-operational ($d=0.33$), age 7-11, concrete operational ($d=0.45$), 12+, formal operational ($d=0.16$).</p> <p><i>Perceptual richness of the manipulatives:</i> Retention ($d=0.28$ rich, $d=0.77$ bland), Problem-solving ($d=-0.27$ rich, $d=0.80$ bland), Transfer ($d=0.48$ rich, $d=-0.02$ bland).</p> <p><i>Mathematical topics:</i> $d=0.21$, $k=10$, Algebra; $d=0.27$, $k=24$) Arithmetic; $d=0.69$, $k=12$, Fractions; $d=0.37$, $k=6$, Geometry; $d=0.58$, $k=3$, Place value.</p>
Holmes (2013)	0.22, 95% CI [0.05, 0.39]	24	<p>ES (d) reported for manipulatives compared to non-use studies ($k=14$).</p> <p>ES (d) for virtual manipulatives compared to physical: 0.20 [-0.05, 0.45] n.s., $k=7$).</p>
Domino (2010): primary	0.39, 95% CI [0.21, 0.56]	24	<p>ES (d) reported for 24 studies with both pre- and post-test measures. Years 4, 5 and 6 (i.e., KS2) appear to benefit most from physical manipulatives (although only Y7 from secondary phase.)</p>
LeNoir (1989)	-	45	<p>LeNoir identified considerable variation in the data with significant & homogeneous effects for acquisition of measurement at Grades K-5 (i.e., primary) of $d=0.24$ and at Grades 6-9 (i.e., KS3 plus Y10) of $d=.43$, but various effects for geometry and place value were either not significantly different from 0 (or the effects were found to be too heterogeneous to report).</p>
Sowell (1989)	0.29	10	<p>ES (d) reported for 10 studies examining the acquisition of broadly stated objectives at Y2-5 when using manipulatives, compared to abstract/symbolic instruction.</p>

			d=0.09, n.s. k=16 for achievement of specific objectives, Y2-9. Significant differences were not found for pictorial versus abstract or concrete versus pictorial. Various other comparisons (attitudes, retention, transfer and a range of years/grades) did not produce clear results, either because of heterogeneity or a very small number of original studies.
--	--	--	--

References

Meta-analyses included

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*(2), 380.

Holmes, A. B. (2013). Effects of Manipulative Use on PK-12 Mathematics Achievement: A Meta-Analysis. Society for Research on Educational Effectiveness.

Domino, J. (2010). The Effects of Physical Manipulatives on Achievement in Mathematics in Grades K-6: A Meta-Analysis. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

LeNoir, P. (1989). The effects of manipulatives in mathematics instruction in grades K-college: A meta-analysis of thirty years of research. Doctoral Thesis, North Carolina State University at Raleigh, ProQuest UMI 8918109.

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education, 20*, 498-505.

Meta-analyses excluded [and reason]

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *NASSP Bulletin, 89*(642), 24-46.
[This meta-analysis addresses the teaching of algebra specifically. However, 3 of 5 original studies categorised as using manipulatives are included in Carbonneau et al.'s (2103) analysis, which also addresses algebra through moderator analysis.]

Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A Meta-Analysis of Algebra Interventions for Learners with Disabilities and Struggling Learners. *Journal of the International Association of Special Education, 15*(1).
[This is focused on algebra for students with learning disabilities and includes only 2 relevant studies.]

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of Instructional Improvement in Algebra: A Systematic Review and Meta-Analysis. *Review of Educational Research, 80*(3), 372-400.
[This meta-analysis addresses the teaching of algebra specifically. However, 3 of 4 original studies categorised as using manipulatives are included in

Carbonneau et al.'s (2103) analysis, which also addresses algebra through moderator analysis.]

Research syntheses

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198. doi: 10.1016/j.learninstruc.2006.03.001

Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning*. London: Nuffield Foundation.

Siegler, R. S., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., . . . Wray, J. (2010). *Developing effective fractions instruction for kindergarten through 8th grade: A practice guide (NCEE #2010-4039)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Other references

Griffiths, R., Back, J., & Gifford, S. (2016). *Making numbers: Using manipulatives to teach arithmetic*. Oxford: Oxford University Press.

Swan, M. (2005). *Improving learning in mathematics*. London: Department for Education and Skills.

7.4 Tasks

What is the evidence regarding the effectiveness of mathematics tasks?

The current state of research on mathematics tasks is more directly applicable to curriculum designers than to schools. Tasks frame, but do not determine, the mathematics that students will engage in, and should be selected to suit the desired learning intentions. However, as with textbooks, *how* teachers use tasks with students is more important in determining their effectiveness. More research is needed on how to communicate the critical pedagogic features of tasks so as to enable teachers to make best use of them in the classroom.

Strength of evidence: LOW

Findings

A classroom mathematics *task* refers to whatever prompt is given to students to indicate what they are to do. This is often distinguished from the *activity* which results from a particular prompt (Christiansen & Walther, 1986), although it is generally acknowledged that it can be difficult to separate a task from the activity that results from it (Watson & Mason, 2007). Tasks are critical to the learning of mathematics, because the tasks used in the mathematics classroom largely define what happens there (Sullivan, Clarke, & Clarke, 2013), as well as contributing to students' perceptions of the nature of mathematics itself. However, how a task is used with students is likely to be more important than the specific details of the task itself (Stein, Remillard, & Smith, 2007).

There is a wealth of literature about mathematics tasks, which is often generously illustrated with examples. Exemplification is critical to communicating task types to teachers, since different teachers interpret task descriptors, such as “rich”, differently (Foster & Inglis, 2017). This suggests that unless curriculum designers give examples of the kinds of tasks intended by a word such as “rich”, their goals are likely to be frustrated, as teachers will interpret the term in different ways. Further research is needed on how to communicate the pedagogic features of tasks in ways that enable teachers to use them effectively in the classroom.

Ahmed (1987, p. 20) listed 10 desirable features of a “rich mathematical activity”, including accessibility, extendibility, potential for surprise, enjoyment and originality, and opportunities for students to pose questions, discuss, make decisions, speculate, make hypotheses and prove (see also Swan, 2008). Swan's (2006) tasks focus on conceptual understanding and frequently address misconceptions directly, within a formative assessment framework. Watson and Mason (2005) designed tasks that exploit variation (Mun Ling, & Marton, 2011) and provide opportunities for students to generate examples of mathematical objects so as to make use of and develop their mathematical powers (Mason, & Johnston-Wilder, 2006). Tasks which invite students to create examples and non-examples can be particularly helpful in broadening and enriching students' example spaces and focusing attention on relevant features of mathematical objects and structure (Watson & Mason, 2005).

In many cases, high-quality mathematics tasks pose a problem for students to solve which admits of multiple solutions or solution approaches that have different levels of mathematical sophistication, commensurate with the capabilities of the students (Ruthven, 2015, p. 314). Inquiry-based, problem-solving tasks are linked in large-scale US empirical studies to significant gains in attainment (e.g., Thomas & Senk,

2001). Sullivan, Clarke and Clarke (2013, p. 57) described what they termed “content-specific open-ended” tasks, which are “accessible by students, able to be used readily by teachers, foster a range of mathematical actions, and contribute to some of the important goals of learning mathematics”. A balance of different kinds of tasks is likely to be desirable, but this can be difficult to achieve if teachers rely excessively on textbooks which are dominated by short, closed exercises.

It is important to note that a rich mathematics task by itself will not automatically produce the intended learning; how the teacher enacts the task is critical (e.g., see Stein, Remillard, & Smith, 2007). Only the teacher who knows their particular students can take account of prior student knowledge and judge how to support and motivate their students to learn mathematics through use of the task. However, Stein, Grover and Henningsen (1996) found that teachers tended to reduce the degree of challenge of tasks, which could be problematic if tasks became mathematically trivialised.

The context used (if any) in a mathematics task is an important factor to consider. Contexts can be distracting and confusing for students (Lubienski, 2000), particularly for low SES students (Cooper and Dunne, 2000). Sometimes contexts are presented illustratively or humorously, but if contexts are supposed to be taken seriously by the students then they should be appropriately realistic, perhaps even relating to topics likely to be of interest or importance to students. The extra cognitive load provided by setting some mathematics within a real-life context may make the task too demanding. Alternatively, a familiar context may help students to appreciate more concretely the mathematical structure lying behind a problem. The Realistic Mathematics Education (RME) programme (De Lange, 1996; Van den Heuvel-Panhuizen, & Drijvers, 2014) uses context not as an add-on to motivate students but to provide realisable/imaginable situations in which students can develop their mathematical understanding.

Anthony and Walshaw (2007) summarised their systematic review by commenting that

The research provides evidence that tasks vary in nature and purpose, with a range of positive learning outcomes associated with problem-based tasks, modelling tasks, and mathematics context tasks. But whatever their format, effective tasks are those that afford opportunities for students to investigate mathematical structure, to generalise, and to exemplify. (p. 140)

It is likely that many tasks, even apparently routine ones, could fulfil these objectives if handled sensitively by a skilful teacher. This suggests that emphasis should be placed on teacher professional development relating to the effective use of a variety of mathematics tasks.

Evidence base

The quantity of research in mathematics task design has increased considerably in recent years, as illustrated by the creation of the International Society for Design and Development in Education (ISDDE) and its journal, *Educational Designer*. Although, as one might expect, we found no experimental studies on task design (only studies on designed interventions), there are many studies concerned with task design. These frequently set out a collection of task design principles, but one difficulty is to decide what constitutes a desirable set of principles. As expected, there are also no meta-analyses of mathematics tasks and just one systematic review which contained

a relevant chapter (Anthony, & Walshaw, 2007). Watson and Ohtani (2015), based on the ICMI Study 22, is an authoritative survey of the current state of the field.

Because of the English language limitation, we have not been able to include the Russian experience, in which task design is central in mathematics teacher education. Nor have we been able to adequately take account of design principles of variation, as applied in Shanghai (and, to some extent, the rest of China), in which effectiveness is discussed deeply and known about but not reported as research. Related to this is also the long tradition of development over decades of problem tasks in Japan, meaning that children now do the same problem tasks as their parents and teachers did when they were at school.

Directness

The variation in context of the various studies examined does not seem a likely threat to the directness of these findings.

Future research

There is a need for more cross-disciplinary research investigating how tasks can be designed in the light of research evidence on how students learn mathematics. We also need to know how to communicate the key features of a task, and the pedagogic opportunities that it offers, to teachers.

References

Meta-analyses included

None

Systematic reviews included

Anthony, G., & Walshaw, M. (2007). *Best evidence synthesis: Effective pedagogy in Pangarau/Mathematics*. Wellington, NZ: Ministry of Education.

Other references

Ahmed, A. (1987). *Better mathematics: A curriculum development study based on the low attainers in mathematics project*. London: HM Stationery Office.

Christiansen, B., & Walther, G. (1986). Task and activity. In B. Christiansen, A. G. Howson, & M. Otte (Eds.), *Perspectives in mathematics education* (pp. 243–307). Dordrecht: Reidel.

Cooper, B., & Dunne, M. (2000). *Assessing children's mathematical knowledge: Social class, sex and problem-solving*. Buckingham: Open University Press.

De Lange, J. (1996). Using and applying mathematics in education. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 49–98). Dordrecht, The Netherlands: Kluwer.

Foster, C., & Inglis, M. (2017). Teachers' appraisals of adjectives relating to mathematics tasks. *Educational Studies in Mathematics*, 95(3), 283–301. <https://doi.org/10.1007/s10649-017-9750-y>

Lubienski, S. (2000). Problem solving as a means toward mathematics for all: An exploratory look through a class lens. *Journal for Research in Mathematics Education*, 31(4), 454–482.

- Mason, J., & Johnston-Wilder, S. (2006). *Designing and using mathematical tasks*. Tarquin Publications.
- Mun Ling, L., & Marton, F. (2011). Towards a science of the art of teaching: Using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies*, 1(1), 7-22.
- Ruthven, K. (2015). Taking design to task: A critical appreciation. In A. Watson, & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 311-320). Heidelberg: Springer.
- Stein, M. (2001). Teaching and learning mathematics: How instruction can foster the knowing and understanding of number. In J. Brophy (Ed.), *Subject-specific instructional methods and activities* (Vol. 8, pp. 114-144). Amsterdam: JAI.
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp 319-369). Charlotte, NC: Information Age Publishing).
- Sullivan, P., Clarke, D., & Clarke, B. (2013). *Teaching with tasks for effective mathematics learning*. New York: Springer.
- Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices*. London: National Institute for Advanced and Continuing Education (NIACE).
- Swan, M. (2008). A Designer Speaks. *Educational Designer*, 1(1). Retrieved from http://www.educationaldesigner.org/ed/volume1/issue1/article3/#s_2 on 11 May 2017.
- Thompson, S., & Senk, S. (2001). The effects of curriculum on achievement in second-year algebra: The example of the University of Chicago School Mathematics Project. *Journal for Research in Mathematics Education*, 32(1), 58-84.
- Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In *Encyclopedia of mathematics education* (pp. 521-525). Springer Netherlands.
- Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: The role of learner generated examples*. Mahweh, NJ: Erlbaum.
- Watson, A., & Mason, J. (2007). Taken-as-shared: A review of common assumptions about mathematical tasks in teacher education. *Journal of Mathematics Teacher Education*, 10, 205–215.
- Watson, A., & Ohtani, M. (Eds.). (2015). *Task design in mathematics education: An ICMI study 22*. Heidelberg: Springer.

7.5 Textbooks

What is the evidence regarding the effectiveness of textbooks?

The effect on student mathematical attainment of using one textbook scheme rather than another is very small, although the choice of a textbook will have an impact on what, when and how mathematics is taught. However, in terms of increasing mathematical attainment, it is more important to focus on professional development and instructional differences rather than on curriculum differences. The organisation of the mathematics classroom and how textbooks can enable teachers to develop students' understanding of, engagement in and motivation for mathematics is of greater significance than the choice of one particular textbook rather than another.

Strength of evidence: HIGH

Findings

Textbooks can play a variety of different roles in the mathematics classroom. At one extreme, they can be viewed as one resource among many, to be dipped into from time to time and drawn from as appropriate within a broader scheme of work. At the other extreme, a textbook may be adopted in a wholesale manner as the basis for the entire mathematics curriculum. In this case, the contents of the textbook (and accompanying teacher guide) can come to define the mathematics to be taught and provide an organised sequence of topics for teachers to use to pace and structure their teaching. If adopted in this way, textbooks can encourage particular pedagogies and teaching strategies and indicate the amount of weight that should be given to different topics, as well as to different aspects of learning, such as routine practice (Howson, 2013).

In their two meta-analyses, Slavin, Lake, & Groff (2007a, 2007b) searched for high-quality studies on elementary and middle school mathematics curricula, and divided the curricula that they examined into three categories:

- reform: NCTM Standards-based NSF-funded curricula stressing “problem solving, manipulatives, and concept development, and a relative de-emphasis on algorithms” (Slavin, Lake, & Groff, 2007b, p. 11), such as *Everyday Mathematics* at elementary level and the *University of Chicago School Mathematics Project (UCSMP)*, *Connected Mathematics*, and *Core-Plus Mathematics* at middle school level;

- traditional, commercial textbooks, which were based on the NCTM Standards but with “a more traditional balance between algorithms, concepts, and problem solving” (Slavin, Lake, & Groff, 2007b, p. 8), such as *McDougal-Littell* and *Prentice Hall*;

- back-to-basics: *Saxon Math*, a “curriculum that emphasizes building students’ confidence and skill in computations and word problems” (Slavin, Lake, & Groff, 2007b, p. 11).

For the elementary school textbooks, Slavin, Lake and Groff (2007a) found a median effect size across the three types of only 0.10 ($k = 13$), even though many of the studies included had methodological problems that might have been expected to inflate the effect sizes. They concluded that “there is limited high-quality evidence supporting differential effects of different math curricula” (p. 17). For the middle and high school textbooks, they found an even smaller overall effect size for mathematics

curricula ($ES = 0.03$, $k = 40$), and outcomes were similar for disadvantaged and non-disadvantaged students and for students of different ethnicities.

Both meta-analyses concluded that there is a “lack of evidence that it matters very much which textbook schools choose” (Slavin, Lake, & Groff, 2007b, p. 44) and that “curriculum differences appear to be less consequential than instructional differences” (Slavin, Lake, & Groff, 2007b, p. 45). They commented that interventions addressing everyday teaching practices and student interactions have more promise than those emphasizing textbooks alone and advise schools to “focus more on how mathematics is taught, rather than expecting that choosing one or another textbook by itself will move their students forward” (Slavin, Lake, & Groff, 2007a, p. 39). The studies used in these meta-analyses cover a diverse range of settings, and there was no clear pattern of any difference in ESs for students according to SES: “Programs found to be effective with any subgroup tend to be effective with all groups” (Slavin, Lake, & Groff, 2007b).

However, the findings from these two US meta-analyses need to be interpreted cautiously for the English context. In most cases, the studies examined compared textbook use to business as usual, which means that some of the control groups also used textbooks, at least for some of the time. Even more importantly, the US does not have a national curriculum, as England does. This means that textbooks may come to define the curriculum in the US to a much greater extent than in England – indeed, as above, a textbook is often described as “a curriculum” in the US.

In schools in England, the balance tends to lie away from the wholesale adoption of textbooks and towards their more selective use. Mathematics teachers in England have consistently made much less use of textbooks than have teachers in other countries. The *TIMSS 2011* study (Mullis et al., 2012) reported that only 29% of mathematics teachers in England used textbooks “as the basis for instruction” at Grade 8 (equivalent to Year 9 in England) (compared to a 77% international average). At Grade 4 (Year 5 in England) the corresponding figure was 10% (compared to a 75% international average).⁷ In each case, these were the second-lowest uses in all the systems surveyed.

Askew, Hodgen, Hossain and Bretscher (2010) found that countries that perform consistently well in international comparative mathematics assessments tend to use more carefully constructed textbooks as the main teaching resource, whereas current textbooks in England tend to be less mathematically coherent and are focused on routine examples (see also, Hodgen, Küchemann & Brown, 2013). Fan, Zhu, & Miao (2013) pointed to many aspects of variation among textbooks from different education systems, both in presentation and in pedagogical structure. English mathematics textbooks are notable for their undemanding routine exercises and fragmented approach, and there has been much criticism of the routine and shallow nature of a great deal of typical English mathematics textbook content. In this connection, Howson (2013) stressed the importance of research focusing on the exercises in textbooks and examining whether they go beyond the routine. In their study of textbooks, Haggarty and Pepin (2002) found that textbooks in England were characterised by unrelated rules and facts aimed at the development of “fluency in

⁷We note that it is possible that the use of textbooks in primary may have increased due to a recent national initiative promoting the use of textbooks, although up to date information is not available: <http://www.mathshubs.org.uk/what-maths-hubs-are-doing/teaching-for-mastery/textbooks/>

the use of routine skills through repeated practice in exercises” (p. 587). There was only “a superficial veneer of including process skills” (p. 586). Newton & Newton (2007) also found that textbooks aimed at primary children in England focused on practising algorithms rather than reasoning and understanding. Continental textbooks tend to have a more intensive focus on fewer ideas, whereas textbooks in England tend to switch topics frequently and revisit them repeatedly (Bierhoff, 1996).

Fan, Zhu and Miao (2013) found that there had been a “general decline both in the amount of material demanding student involvement and in the percentage of that material requiring higher-order thinking” (p. 638). They also found that problem-solving tasks were simplistic, opportunities for deductive reasoning were largely absent and the majority of problems had no connection with the real world. They also stressed the critical role that teachers play in determining how they use textbooks and, in particular, what they choose to omit. Important mathematical connections were often not explicitly made in textbooks. For example, Levin (1998) found that in US elementary, middle school, and algebra textbooks, fractions and division were generally presented separately rather than in ways that contributed to building meaningful connections.

It seems clear that although textbooks are important, simply providing “better” textbooks will not by itself improve learning. Teachers have much greater effects on student attainment than textbooks or other resources, so textbooks need to be seen as part of a programme of change that includes professional development (PD); indeed, good textbooks might be enablers of this. The closest thing in England in recent years to wholesale adoption of a single textbook scheme is the National Numeracy Strategy (DfEE, 1999; DfEE, 2001), where the Framework comprised something closer to a curriculum than to a textbook, with pedagogical advice and a considerable range and variety of examples of tasks. In the primary phase, at least, the Strategy appears to have had a large system-wide effect of about 0.18 (see Brown et al., 2003), and it is noteworthy that the NNS was partially research-based (Brown, et al., 1998) and enjoyed PD, external support and headteacher engagement. The importance of these factors should not be underestimated.

Evidence base

In recent decades there has been a large increase in the amount of research on mathematics textbooks, a subject which had previously been relatively neglected (Fan, Zhu, & Miao, 2013; Howson, 2013). We found one recent systematic review (Fan, Zhu, & Miao, 2013) and two meta-analyses: one focused on elementary schools (Slavin, Lake, & Groff, 2007a) and the other on middle and high school (Slavin, Lake, & Groff, 2007b).

Fan, Zhu, & Miao (2013) carried out a systematic search of literature published over the last 60 years. The authors noted that most of the studies that they found were small-scale exploratory studies by individual researchers, which generally focused on textbook use by teachers, rather than by students.

Slavin, Lake, & Groff (2007a, 2007b) included only randomized or matched control group studies in which the two groups were equal at pre-test and the intervention lasted at least 12 weeks. A minimum treatment duration of 12 weeks was required in order to focus on practical programmes intended for use across a whole school year.

Directness

In the US, where the majority of the studies on mathematics textbooks have been carried out, textbook use in mathematics is greater than in England, and textbooks are frequently referred to as “mathematics curricula”. As described above, in England the predominant approach appears to be sourcing material for lessons from a diverse selection of books and websites. While this approach could lead to some higher-quality lessons than those offered in any single textbook, it is time consuming for schools and makes coherence and balance harder to attain. This approach, coupled with frequent changes to the National Curriculum, may also have made it harder for publishers to fund the development of high-quality textbooks. We note that recent initiatives to promote textbooks inspired by those used in Singapore and Shanghai may affect the use of textbooks in English primary schools, but these initiatives have yet to be rigorously evaluated.

This is not to say that all textbooks are alike. Fan, Zhu, & Miao (2013) commented that “remarkable differences were found in textbooks from different series and particularly from different countries, which seems to [them] to point not only to the lack of consensus in textbook development, but also to the inseparability of textbooks from the cultural and social background.” (p. 640) The choice of one particular textbook over another will have implications for what, when and how mathematics is taught. Hence, schools and teachers do need to give careful consideration to textbook choice, and guidance should be provided, but choice of textbook by itself is unlikely to raise attainment in mathematics.

We were not able to find meta-analyses specifically looking at the use of ebooks in the classroom.

Threat to directness	Directness (1 low – 3 high)	Notes
Where and when the studies were carried out	2	Differences in textbook use in the US and England reduce the directness of these findings.
How the intervention was defined and operationalised	2	Uncertainty over the extent to which textbooks were adopted in their entirety.
Any reasons for possible ES inflation	2	Concern regarding attrition of schools in post-hoc analyses and lack of clear controls. The fact that the counterfactual sometimes included textbook use is problematic.
Any focus on particular topic areas	3	
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Quality	Study inclusion dates	Comments

Slavin, Lake, & Groff (2007b)	0.03	40	3	1971-2008	<p>Looked at research on the achievement outcomes of mathematics programmes for middle and high schools.</p> <p>Effect sizes were somewhat higher for the <i>Saxon</i> textbooks (weighted mean ES=0.14 in 11 studies) than for the NSF-supported textbooks (median ES=0.00 in 26 studies). However, the NSF programmes add objectives not covered in traditional texts, so to the degree to which those objectives are seen as valuable, these programmes are adding impacts not registered on the assessments of content covered in all treatments. Among 3 studies of traditional mathematics curricula, one (<i>Prentice Hall Course 2</i>) found substantial positive effects, but two found no differences.</p> <p>The weighted mean effect size for 24 studies of NSF-funded programs was 0.00, even lower than the median of +0.12 reported for elementary NSF-funded programs.</p> <p>It has been suggested that possible misalignment between the NSF-sponsored curricula and the standardized tests used to measure their effectiveness could account for these small effect sizes, but Slavin et al. do not think this a likely explanation.</p>
Slavin, Lake, & Groff (2007a)	0.10	13	3		<p>Most of the studies comparing mathematics curricula are of "marginal methodological quality": "Ten of the 13 qualifying studies used post-hoc matched designs in which control schools, classes, or students were matched with experimental groups after outcomes were known. Even though such</p>

					<p>studies are likely to overstate program outcomes, the outcomes reported in these studies are modest. The median effect size was only +0.10. The enormous ARC study found an average effect size of only +0.10 for the three most widely used of the NSF-supported mathematics curricula, taken together. Riordan & Noyce (2001), in a post-hoc study of <i>Everyday Mathematics</i>, did find substantial positive effects (ES=+0.34) in comparison to controls for schools that had used the program for 4-6 years, but effects for schools that used the program for 2-3 years were much smaller (ES=+0.15). This finding may suggest that schools need to implement this program for 4-6 years to see a meaningful benefit, but the difference in outcomes may just be a selection artifact, due to the fact that schools that were not succeeding may have dropped the program before their fourth year. The evidence for impacts of all of the curricula on standardized tests is thin. The median effect size across five studies of the NSF-supported curricula is only +0.12, very similar to the findings of the ARC study.”</p>
--	--	--	--	--	--

References

Meta-analyses included

- Slavin, R. E., Lake, C., & Groff, C. (2007a). *A Best-Evidence Synthesis Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis*. Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University.
- Slavin, R. E., Lake, C., & Groff, C. (2007b). *Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis*. Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University.

Meta-analyses excluded

- Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79(2), 839-911.
- Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research*, 78(3), 427-515.

[Shortened versions of 2007 ones, so original 2007 reports used.] **Systematic reviews included**

- Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: development status and directions. *ZDM Mathematics Education*, 45(5), 633-646.

Other references

- Askew, M., Hodgen, J., Hossain, S., & Bretscher, N. (2010). *Values and variables: Mathematics education in high-performing countries*. London: Nuffield Foundation.
- Bierhoff, H. (1996). *Laying the foundations of numeracy: A comparison of primary school textbooks in Britain, Germany and Switzerland*. London: National Institute for Economic and Social Research.
- Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385.
- Brown, M., Askew, M., Millett, A., & Rhodes, V. (2003). The key role of educational research in the development and evaluation of the National Numeracy Strategy. *British Educational Research Journal*, 29(5), 655-672.
- DfEE (1999). *The National Numeracy Strategy: Framework for teaching mathematics from Reception to Year 6*. London: Department for Education and Employment.
- DfEE (2001). *Framework for teaching mathematics Years 7, 8 and 9*. London: Department for Education and Employment.
- Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, 28(4), 567-590.
- Hodgen, J., Küchemann, D., & Brown, M. (2010). Textbooks for the teaching of algebra in lower secondary school: are they informed by research? *Pedagogies*, 5(3), 187-201. doi:10.1080/1554480X.2013.739275
- Howson, G. (2013). The development of mathematics textbooks: historical reflections from a personal perspective. *ZDM Mathematics Education*, 45(5), 647-658.
- Levin, S. W. (1998). *Fractions and division: Research conceptualizations, textbook presentations, and student performances* (Doctoral dissertation, University of Chicago, 1998). Dissertation Abstracts International, 59, 1089A.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA / Amsterdam: TIMSS & PIRLS

International Study Center, Lynch School of Education, Boston College /
International Association for the Evaluation of Educational Achievement (IEA).

Newton, D. P., & Newton, L. D. (2007). Could elementary mathematics textbooks help give attention to reasons in the classroom? *Educational Studies in Mathematics*, 64(1), 69-84.

What Works Clearinghouse (2006). *Elementary school math*. Washington, DC: U.S. Department of Education.

What Works Clearinghouse (2007). *Middle school math*. Washington, DC: U.S. Department of Education.

8 Mathematical Topics

8.1 Overview

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of specific topics within mathematics?

The mathematics national curriculum covers a range of topics and strands, including: number, algebra, ratio, proportion and rates of change, geometry and measures, probability, and statistics. Elsewhere in this review, we have examined 'generic' approaches to teaching and learning mathematics, such as the use of concrete manipulatives, which are applicable across these topics and strands. It would be reasonable to assume that, whilst there are many similarities in teaching approaches, there are likely to be some differences. However, we found the evidence base to be limited in two ways. First, as Nunes et al. (2009) observe, there is little research in general on the "technicalities of teaching", or how to teach learners in specific topics. Second, the literature base is skewed. Aside from one meta-analysis relating to the use of dynamic geometry software,⁸ we found no meta-analyses addressing effective approaches to teaching geometry, measures, probability or statistics. Aside from the meta-analyses relating to calculator use, we identified four meta-analyses focused on number and arithmetic/calculation, all concerned with approaches for learners with either learning or other cognitive disabilities or special educational needs.

We identified three meta-analyses concerned with algebra, one of which addresses the particular needs of those with learning disabilities. We also identified three relevant What Works Clearinghouse (WWC) practice guides from the US, one concerned with teaching algebra, one with teaching fractions and another with teaching "struggling" learners. In order to address the gaps in the evidence base, we have drawn additionally on several systematic reviews (e.g., Nunes et al., 2009). These reviews are mainly focused on how children *learn* rather than how to *teach*, although there is a great deal of guidance on what to emphasise in teaching. Hence, we use these to interpret and extend the WWC findings, in particular those relating to the teaching of fractions.

We focus on the four mathematical topics: algebra, number (including calculation and multiplicative reasoning), geometry and measures, and probability and statistics. We note that the research base on the effectiveness of teaching approaches for geometry and measures, and for probability and statistics, is extremely limited.

References

- Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420-464). New York: Macmillan
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

⁸ See Technology module.

8.2 Algebra

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of algebra?

Learners generally find algebra difficult because of its abstract and symbolic nature and because of the underlying structural features, which are difficult to operate with. This is especially the case if learners experience the subject as a collection of arbitrary rules and procedures, which they then misremember or misapply. Learners benefit when attention is given both to procedural and to conceptual teaching approaches, through both explicit teaching and opportunities for problem-based learning. It is particularly helpful to focus on the structure of algebraic representations and, when solving problems, to assist students in choosing deliberately from alternative algebraic strategies. In particular, worked examples can help learners to appreciate algebraic reasoning and different solution approaches.

Strength of evidence: MEDIUM

For the purposes of this review, we define algebra as a powerful set of mathematical tools used to express generalisations and relationships between numbers, expressions, functions and other mathematical objects, using symbols, graphs, numbers and words (see, e.g., Kieran, 2004). There is a great deal of evidence that learners encounter significant difficulties with algebra (Hart, 1984; Hodgen et al., 2012). In common with many researchers, Rakes et al. (2010) argue that this is due to a predominance of drill and practice approaches to teaching that do not facilitate algebraic understanding. They highlight three conceptual challenges in the learning of algebra:

The abstract nature of algebra: In the transition to algebraic thinking, learners are required to think more abstractly; for example, by making generalisations about expressions or equations using rules and logical relations (Nunes et al., 2009). This can require learners to process many pieces of complex information at the same time, thus increasing cognitive load (Star et al., 2015).

The meaning of algebraic symbols: In algebra, letters are used to represent unknown numbers, variables, parameters and constants. There is an extensive literature on learners' difficulties and misconceptions regarding the interpretation of letters, which can prevent learners from connecting the symbols to their meanings (Küchemann, 1981; Nunes et al., 2009).

The structural characteristics of algebra: Algebra involves the study of structures and systems abstracted from number and relations (Kaput, 2008). Without an appreciation of this structure, learners often conceive of algebra as a collection of arbitrary rules and, for example, misapply or misremember rules for manipulating algebraic expressions or equations (Nunes et al., 2009).

By coding the literature, Rakes et al. (2010) identified five categories of approaches to the teaching of algebra that they judge to be distinct from drill and practice. The five categories were: interventions focused on changes to teaching (including both cooperative learning and mastery approaches), concrete manipulatives, non-

technology-based curricula,⁹ technology tools (both software and calculators), and technology-based curricula (mainly computer-aided instruction of various types). In each case, the categorisation was deliberately broad in order to include, and thus compare, the effects of both procedurally and conceptually based approaches. Rakes et al. (2010) found some evidence to support the efficacy of all five approaches. In addition, they found positive effects for both procedurally and conceptually-focused approaches. Whilst this indicates that it is valuable to use procedural and conceptual teaching approaches, it provides limited actionable guidance for teachers on what specific approaches to use, as well as when and how to integrate them.

Haas's (2005) meta-analysis identified from the literature six approaches to teaching algebra: cooperative learning, communication and study skills, explicit teaching,¹⁰ problem-based learning, technology-aided learning, and manipulatives, models and multiple representations. He finds medium-sized effects for direct instruction and problem-based learning ($d=0.55$ and $d=0.52$, respectively), smaller effects for manipulatives and cooperative learning ($d=0.38$ and $d=0.34$, respectively) and near negligible (but positive) effects for communication and technology-based approaches. Haas argued that these findings do not imply that teachers should avoid using communication and study skills approaches or technology (or manipulatives and cooperative learning), but rather he observed that both explicit teaching and problem-based learning can encompass each of these approaches, each of which "represents less an overarching approach to teaching and more a tool to be incorporated within a lesson [and] teachers should possess a wide repertoire of such tools and strategies" (p. 40).

Elsewhere in this review, we provide evidence for the efficacy of explicit teaching as an approach and, specifically for teaching algebra; Hass argues strongly on the basis of his review for greater use of explicit teaching. It is important to note that Hass argues that explicit teaching should *not* be the only approach that teachers adopt, and that teachers need to adapt their approach to changes in the teaching and learning situation so that learners perceive learning as "meaningful and significant" (p. 38). Thus, assessment plays a key role not only in understanding what students know, but also in informing teacher judgments about the most appropriate teaching approaches to address the next steps in learning. However, whilst Haas's meta-analysis provides evidence to warrant greater use of explicit teaching, it does not provide specific guidance on what practitioners should do.

In a What Works Clearinghouse practitioner guide on the teaching of algebra, Star et al. (2015) highlight three evidence-based approaches that provide useful guidance for explicit teaching in algebra, and which place emphasis on both procedural and conceptual understanding:

Use worked examples to enable learners to analyse algebraic reasoning and strategies: Worked examples, or 'solved problems', enable learners to see

⁹ In effect, Rakes et al.'s (2010) non-technology curricula consist of textbook schemes that are commonly used in the US. They include both traditional and reform-based schemes in this category.

¹⁰ Haas (2005) used the term *direct instruction*, which we have categorised in more general terms as *explicit instruction*. He defined direct instruction as follows: "Establishing a direction and rationale for learning by relating new concepts to previous learning, leading students through a specified sequence of instructions based on predetermined steps that introduce and reinforce a concept, and providing students with practice and feedback relative to how well they are doing." (p. 28).

the problem and the solution together. By removing the need to carry out each step in a solution, worked examples reduce cognitive load, thus enabling learners to discuss and analyse the reasoning and strategies involved. Worked examples may be complete, incomplete or incorrect, deliberately containing common errors and misconceptions for learners to uncover.

Teach learners to recognise and use the structure of algebraic representations:

An explicit focus on structure can help learners to “make connections among problems, solution strategies, and representations that may initially appear different but are actually mathematically similar” (Star et al., 2015, p. 16). Teaching should encourage learners to use language that reflects algebraic structure and to notice that different mathematical representations (e.g., symbolic, numeric, verbal or graphical) can communicate, or place different emphasis on, different characteristics of algebraic expressions, equations, relationships or functions. Nunes et al. (2009) recommend that learners “read numerical and algebraic expressions relationally, rather than as instructions to calculate (as in substitution)” (p. ?); the same is also necessary with regard to the equals sign (Jones & Pratt, 2012).

Teach learners to intentionally choose from alternative algebraic strategies when solving problems:

Choosing, comparing and evaluating different strategies can develop learners’ procedural fluency and conceptual understanding. Encouraging learners to compare strategies can enable them to build on their existing knowledge. Teaching should encourage learners to articulate, and justify, the reasoning underlying different strategies.

Whilst Star et al. (2015) consider all three approaches to be evidence-based, they judge the evidence to be stronger for alternative strategies (moderate evidence) than for worked examples and algebraic structure (limited strength). Additionally, the three approaches resonate with many of the findings of Nunes et al.’s (2009) review.

One further meta-analysis examined approaches to algebra teaching for students with learning disabilities (or at risk of developing learning disabilities). Hughes et al. (2014) identified two potentially effective approaches, each with limited evidence: cognitive/model-based approaches using explicit instruction to teach problem-solving strategies, and concrete-pictorial-abstract approaches.

Evidence base

We found three meta-analyses examining the effect of teaching approaches in algebra, one of which is focused on learners with learning disabilities. There is a great deal of overlap between the two remaining meta-analyses. The largest and most recent of these (Rakes et al., 2010) is of high quality and draws on a larger number of original studies.

Meta-analysis	Focus	k	Quality	Date Range	Overlap with Rakes et al.
Rakes et al. (2010)	Teaching methods in algebra (mainly secondary).	82	3	1968-2008	N/A
Haas (2005)	Teaching methods in secondary algebra	26	2	1980-2002	20 (76.9%)

Hughes et al. (2014)	Teaching methods in algebra for learners with disabilities & struggling learners (at risk for a mathematics disability).	12	2	1985-2002	1 (8.3%)
----------------------	--	----	---	-----------	----------

Directness

Our overall judgement is that the available evidence is of high directness.

The majority of the studies examined in these meta-analyses are set in the US and inevitably the studies were designed around the particularities of the US school system, in which learners have an entire year of mathematics labelled as “Algebra”. However, the problems that students encounter in algebra in the US and English systems are very similar (Kieran, 1992; Küchemann, 1981; Nunes et al., 2009). Moreover, the two main meta-analyses (Rakes et al., 2010; Haas, 2005) focus on general approaches that we judge to be largely applicable in both systems. The WWC Practice Guide (Star et al., 2015) is judged to highlight approaches that would be applicable in the English context, because similar approaches are highlighted in Nunes et al.’s (2015) review.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	Most original studies were US based, but results judged to be applicable to England.
How the intervention was defined and operationalised	3	The meta-analyses focus on generic approaches (e.g., direct instruction, use of multiple representations) rather than highly-structured interventions
Any reasons for possible ES inflation	3	Not specifically. The meta-analysis (Hughes et al., 2015) related to the LD population was taken out of the main analysis.
Any focus on particular topic areas	3	NA
Age of participants	3	Mainly secondary, but some upper secondary and college level in Rakes et al. (2010).

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Effect of different teaching approaches on attainment in algebra</i>			
Rakes et al (2010)	0.21 – 0.32	82	Instructional change (including both cooperative learning and mastery approaches): 0.32 (SE 0.030)

			Concrete manipulatives: 0.32 (SE 0.89) Curricula (US textbook schemes): 0.21 (SE 0.024) Technology tools (both software and calculators): 0.30 (SE 0.046) Technology-based curricula (e.g. computer-aided instruction): 0.31 (SE 0.050) Bayes effects reported. Rakes et al. also calculate “design effect adjusted random effects”.
Haas (2005)	0.55 – 0.07	22	Cooperative learning: 0.34 Communication and study skills: 0.07 Direct instruction (explicit teaching): 0.55 Problem-based learning: 0.52 Technology-aided learning: 0.07 Manipulatives, models & multiple representations: 0.38

Comparison of procedurally and conceptually focused approaches to teaching algebra

Rakes et al. (2010)	See comment	82	Rakes et al. report two approaches to the calculation of ESSs: Bayes adjusted fixed effects and design effect adjusted random effects. These result in different relative magnitudes for procedural and conceptual approaches & Rakes et al. argue that this demonstrates the potential greater efficacy of conceptually-based approaches.									
			<table><tr><td></td><td>Bayes</td><td>Design effect adjusted</td></tr><tr><td>Concept</td><td>0.232 (SE 0.023)</td><td>0.467 (SE 0.099)</td></tr><tr><td>Procedur e</td><td>0.301 (SE 0.023)</td><td>0.214 (SE 0.044)</td></tr></table>		Bayes	Design effect adjusted	Concept	0.232 (SE 0.023)	0.467 (SE 0.099)	Procedur e	0.301 (SE 0.023)	0.214 (SE 0.044)
	Bayes	Design effect adjusted										
Concept	0.232 (SE 0.023)	0.467 (SE 0.099)										
Procedur e	0.301 (SE 0.023)	0.214 (SE 0.044)										

Effect of teaching approaches on attainment in algebra for learners with learning disabilities or struggling learners at risk of developing learning disabilities

Hughes et al. (2015)	0.62, 95% CI [0.48, 0.76]	8	Cognitive/model-based approaches using explicit instruction to teach problem-solving strategies: 0.68, 95% CI [0.48, 0.88], $k=4$
----------------------	---------------------------	---	---

			Concrete-pictorial-abstract ¹¹ approaches: 0.52, 95% CI [0.28, 0.76], $k=2$ Insufficient information or too few original studies to calculate aggregated ESs for the effects of co-teaching, graphic organisers, single-sex instruction and technology.
<i>Effective techniques to teaching algebra</i>			
<i>Star et al. (2015) (WWC Practice Guidance)</i>			Uses What Work Clearinghouse standards
Worked examples		4	Minimal evidence base
Algebraic structure		6	Minimal evidence base
Alternative strategies		10	Moderate evidence base

References

Meta-analyses included

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin*, 89(642), 24-46.).

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: a systematic review and meta-analysis. *Review of Educational Research*, 80. doi:10.3102/0034654310374880

Secondary Meta-analyses

Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A Meta-Analysis of Algebra Interventions for Learners with Disabilities and Struggling Learners. *Journal of the International Association of Special Education*, 15(1).

Systematic reviews

Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning*. London: Nuffield Foundation.

Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., Lyskawa, J., McCallum, W. G., Porath, J., & Zbiek, R. M. (2015). *Teaching strategies for improving algebra knowledge in middle and high school students* (NCEE 2014-4333). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: <http://whatworks.ed.gov>

¹¹Hughes et al. (2015) use the term 'concrete-representational-abstract'.

Other references

- Hart, K. (Ed.) (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- Hodgen, J., Brown, M., Küchemann, D., & Coe, R. (2011). *Why have educational standards changed so little over time: The case of school mathematics in England*. Paper presented at the British Educational Research Association (BERA) Annual Conference, Institute of Education, University of London.
- Jones, I., & Pratt, D. (2012). A Substituting Meaning for the Equals Sign in Arithmetic Notating Tasks. *Journal for Research in Mathematics Education*, 43(1), 2-33.
- Kaput, J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5-17). New York, NY: Lawrence Erlbaum.
- Kieran, C. (2004). The core of algebra: Reflections on its main activities. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The future of the teaching and learning of algebra: The 12th ICMI Study* (pp. 21-33). Dordrecht, The Netherlands: Kluwer.
- Küchemann, D. E. (1981). *The understanding of generalised arithmetic (algebra) by secondary school children* (PhD thesis). Chelsea College, University of London.

8.3 Number and calculation

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of number and calculation?

Number and numeric relations are central to mathematics. Teaching should enable learners to develop a range of mental and other calculation methods. Quick and efficient retrieval of number facts is important to future success in mathematics. Fluent recall of procedures is important, but teaching should also help learners understand how the procedures work and when they are useful. Direct, or explicit, teaching can help learners struggling with number and calculation. Learners should be taught that fractions and decimals are numbers and that they extend the number system beyond whole numbers. Number lines should be used as a central representational tool in teaching number, calculation and multiplicative reasoning across Key Stages 2 and 3.

Strength of evidence: MEDIUM

Findings

Our literature search found eight meta-analyses specifically addressing number and calculation, together with a US-focused *What Works Clearinghouse* practitioner guide on the teaching of fractions. Four of the eight meta-analyses were concerned with calculator use and the other four addressed the teaching and learning of children and young people with learning disabilities. We found no meta-analyses specifically addressing the teaching of multiplicative reasoning, number sense, estimation, or general teaching of calculation. Given the importance of these areas and quantitative reasoning (e.g., Hodgen & Marks, 2013), it is surprising that the evidence base relating to the teaching of number is so limited.

There is a great deal of research about how children learn number and calculation in general (Fuson, 1992) and specific to the development of number sense (Sowder, 1992), additive reasoning (e.g., Nunes et al., 2009), multiplicative reasoning (e.g., Behr, et al., 1992; Lamon, 2007), the relationship between number and algebra (e.g., Nunes et al., 2009) and learners' common errors and misconceptions (e.g., Hart, 1981; Ryan & Williams, 2009). A number of implications for teaching arise from this research base. For example, Nunes et al. (2009) indicate that teaching should enable learners to understand the inverse relation between addition and subtraction, to develop multiplicative reasoning alongside additive reasoning, to use their understanding of division situations to understand equivalence and order of fractions, and to understand the equals sign as meaning 'equal to' or 'equivalent to' rather than as an instruction to evaluate something. However, enacting such principles is not straightforward. Specifically, evidence on what teaching approaches and interventions teachers can use (or on what other outcomes should be given a lower priority in order to achieve these learning outcomes) is weak.

Developing calculation and fluency with number

It is instructive to consider the research base on calculator use, which we summarised in a separate module (see Calculator module). The meta-analyses are based on an extensive set of original studies. Broadly, this research indicates that calculators can be a useful pedagogic tool if integrated into the teaching of calculation more generally, and specifically the teaching of mental methods. Hence,

taken together with the additional evidence cited in the Calculator module, this suggests the following recommendation:

Teach learners to use a range of mental and other calculation methods. Help learners to regulate their use of calculators to complement mental methods.

However, calculators are a tool and, whilst important, form only one element of an integrated approach to the teaching of calculation. The four meta-analyses on calculators provide only limited guidance on the specifics of such an integrated approach to the teaching of calculation. Indeed, much of what constitutes 'best practice' in the teaching of calculation is based largely on inferences from research on how learners learn, rather than on specific evidence on teaching approaches.

So, for example, Thompson (2001) criticises the teaching approach described in the National Numeracy Strategy's *Framework for Teaching Mathematics* (DfEE, 1999) as follows:

The Framework also describes a clear teaching progression for calculation, starting from mental methods, passing through jottings, informal written methods, formal algorithms using expanded notation, and culminating in the learning of standard algorithms. Research is urgently needed to ascertain the extent to which this seemingly logical progression is pedagogically sound. (p. 18)

We note that our literature search was largely focused on identifying meta-analyses, and it may be that a sufficiently large set of rigorously designed studies does in fact exist, but has yet to be synthesised. Hence, there is an urgent need to conduct a review of this literature to ascertain whether a meta-analysis is possible and to establish what additional research is needed in order to understand how to teach calculation.¹²

Supporting learners struggling with number and calculation

We identified one relevant meta-analysis (Kroesbergen & Van Luit, 2003), focused on students with special educational needs, and we additionally draw on a US-focused What Works Clearinghouse (WWC) practitioner guide on the teaching of students struggling with mathematics.

Kroesbergen & Van Luit (2003) synthesised 58 studies reporting interventions targeted at low-performing students, students with learning difficulties and those with "mild mental retardation". Most studies were focused on basic facts ($d=1.14$, $k=31$, $N=1324$) as opposed to preparatory arithmetic ($d=.92$, $k=13$, $N=664$) and problem solving ($d=.63$, $k=17$, $N=521$). Separate meta-analyses were conducted for each of these and the effect sizes were found to be heterogeneous in each case. For basic facts, the variance was explained by study design, peer-tutoring (which was found to be less effective than not), age (interventions for older students were more effective) and instruction method (direct instruction [DI] more effective than self-instruction or mediated instruction). Overall, self-instruction ($d=1.45$) produced a larger ES than DI ($d=.94$) or mediated instruction ($d=.34$). In other words, self-instruction, providing a set of verbal prompts, is more effective in general than DI, but DI appears to be more effective for learning basic facts (at least for students with SEN). The authors

¹² We note that a systematic review of interventions in primary mathematics is currently being conducted by Victoria Sims, Camilla Gilmore and Seaneen Sloane and is due to report in 2018: <http://www.nuffieldfoundation.org/review-interventions-improve-primary-school-maths-achievement>

compared instruction by teacher ($d=1.05$) and by computer ($d=.51$), arguing that, whilst a computer can be very helpful, it cannot replace instruction by a teacher.

Gersten, Beckman et al.'s (2009) *What Works Clearinghouse* (WWC) practitioner guide focuses on “assisting students struggling with mathematics... [in] elementary and middle schools”. Four of the eight recommendations are particularly relevant to the teaching of calculation.¹³ The focus of these recommendations is on interventions; however, we consider these recommendations to be relevant to many learners:

Teaching during the intervention should be explicit and systematic. The guidance highlights the effectiveness of “direct, teacher-guided, explicit instruction” (see also, NMAP, 2008), which they recommend should include both “easy and hard” problems, guided practice, and specific feedback. Teachers should make their approach explicit by thinking aloud when modelling strategies and methods.

Provide learners with opportunities to solve word problems with similar mathematical structures. The guidance highlights the value of using well-chosen problems to “give meaning to mathematical operations such as subtraction or multiplication” (p. 26) by using representations such as the bar model.

Help learners to use visual representations of mathematical ideas. (See manipulatives and representations module).

Provide dedicated time of “about 10 minutes” during each intervention session to build fluent retrieval of arithmetic facts. The guidance highlights the importance of providing learners with regular, structured opportunities to practise ideas previously covered in depth, and emphasises the importance of derived facts.

Fractions, decimals and proportional reasoning

As already noted, we did identify a *What Works Clearinghouse* (WWC) practitioner guide on “effective fractions instruction for kindergarten through 8th grade” (Siegler et al., 2010). Aside from the WWC guidance referred to above on helping students struggling with mathematics (Gersten, Beckman et al., 2009), this is one of only three WWC guides that focus on the specifics of teaching particular mathematical topics, and we refer to the other WWC practitioner guides in the module on algebra and the module on problem-solving. The title of this guidance reflects the importance accorded to fractions within the US curriculum, although the focus on fractions in the title of this one is somewhat misleading. Siegler et al. emphasise links between fractions and proportional reasoning more generally, and fractions is taken here to include decimals, as well as how fractions may be used to express multiplicative relations, including percentages and the relationship between division and fractions (Nunes et al., 2009).

¹³Gersten, Beckman et al.'s (2009) remaining recommendations cover screening to assess the need for intervention, the focus of interventions (whole numbers for KS2, and rational numbers for KS2 and 3), monitoring progress and including motivational strategies. Screening is judged to be supported by a moderate level of evidence, whilst the other three are judged to be supported by a low level of evidence.

Four of Siegler et al.'s five recommendations apply to teaching approaches and are framed in ways that are actionable in the classroom.¹⁴ Reflecting the narrow focus of the title, the recommendations refer almost exclusively to fractions, and we have consequently reworded these to better frame them for the context of school mathematics in England.

1. Build on learners' informal understanding of sharing and proportionality to develop early fraction and division concepts.
2. Teach learners that fractions and decimals are numbers and that they expand the number system beyond whole numbers. Use number lines as a central representational tool in teaching number, calculation and multiplicative reasoning across Key Stages 2 and 3.¹⁵
3. Teach learners to understand procedures for computations with fractions, decimals and percentages.
4. Develop learners' conceptual understanding of strategies for solving ratio, rate, and proportion problems before exposing them to cross-multiplication as a procedure to solve such problems.

Whilst Siegler et al. (2010) consider all four approaches to be evidence-based, they judge the evidence base to be stronger for their recommendations similar to our 2 and 3 (moderate evidence) than for their recommendations similar to our 1 and 4 (limited strength). Additionally, the four approaches resonate very strongly with the findings of Nunes et al.'s (2009) review.

Evidence base

As discussed above, the evidence base is very limited. See Calculator module for quality judgments, effects sizes and other details of the meta-analyses concerned with calculator use.

Directness

Our overall judgement is that the available evidence is of high directness, although the evidence base is patchy and limited.

Despite differences in the US and English curricula, the WWC Practice Guide (Siegler et al., 2010) is judged to highlight approaches that would be applicable in the English context, because similar approaches are highlighted in Nunes et al.'s (2015) review.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	Most original studies were US based, which places greater emphasis on fractions than is the case in England. Nevertheless, results judged to be applicable to England. There are very few original studies.

¹⁴The fifth recommendation addresses the professional development of teachers: *Professional development programs should place a high priority on improving teachers' understanding of fractions and of how to teach them.*

¹⁵The Singapore bar method used in many schools in England is a valuable and pedagogically useful form of the number line that is relatively concrete (see Ng & Lee, 2009, for a discussion). It is valuable to help learners to build on such models to develop more general number line representations.

How the intervention was defined and operationalised	3	The practitioner guidance focuses on generic approaches (e.g., direct instruction, use of multiple representations) rather than highly-structured interventions
Any reasons for possible ES inflation	-	No effect sizes reported.
Any focus on particular topic areas	3	Focused on fractions
Age of participants	3	

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Quality	Comment
<i>Effect of interventions for students struggling with mathematics</i>				
Kroesberg en & Van Luit (2003)	1.14 (basic facts)	58	2	Basic facts ($d=1.14$, $k=31$, $N=1324$) Preparatory arithmetic ($d=.92$, $k=13$, $N=664$) Problem solving ($d=.63$, $k=17$, $N=521$).

Systematic review	No of studies (k)	Comment
<i>Assisting students struggling with mathematics</i>		
Gersten et al. (2009). (WWC Practice Guidance)		Uses What Work Clearinghouse standards
1. Screen all students to identify those at risk for potential mathematics difficulties and provide interventions to students identified as at risk.	N/A	Moderate evidence
2. Instructional materials for students receiving interventions should focus intensely on in-depth treatment of whole numbers in kindergarten through grade 5 and on rational numbers in grades 4 through 8. These materials should be selected by committee.	3	Low evidence
3. Instruction during the intervention should be explicit and systematic. This includes providing models of proficient problem	6	Strong evidence

solving, verbalization of thought processes, guided practice, corrective feedback, and frequent cumulative review.		
4. Interventions should include instruction on solving word problems that is based on common underlying structures.	9	Strong evidence
5. Intervention materials should include opportunities for students to work with visual representations of mathematical ideas and interventionists should be proficient in the use of visual representations of mathematical ideas.	13	Moderate evidence
6. Interventions at all grade levels should devote about 10 minutes in each session to building fluent retrieval of basic arithmetic facts.	7	Moderate evidence
7. Monitor the progress of students receiving supplemental instruction and other students who are at risk.	N/A	Low evidence
8. Include motivational strategies in [...] interventions.	2	Low evidence
<i>Fractions, decimals and proportional reasoning</i>		
Siegler et al. (2010). <i>Developing effective fractions instruction for kindergarten through 8th grade</i> . (WWC Practice Guidance)		Uses What Work Clearinghouse standards
Build on students' informal understanding of sharing and proportionality to develop initial fraction concepts	9	Minimal evidence base
Help students recognise that fractions are numbers and that they expand the number system beyond whole numbers. Use number lines as a central representational tool in teaching this and other fraction concepts from the early grades onward	9	Moderate evidence base
Help students understand why procedures for computations with fractions make sense.	7	Moderate evidence base
Develop students' conceptual understanding of strategies for solving ratio, rate, and proportion problems before exposing them to cross-multiplication as a procedure to use to solve such problems.	6	Minimal evidence base
Professional development programs should place a high priority on improving teachers'	4	Minimal evidence base

understanding of fractions and of how to teach them.		
--	--	--

References

Meta-analyses included

Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and special education, 24*(2), 97-114.

Meta-analyses excluded

Browder, D. M., Spooner, F., Ahlgrim-Dezell, L., Harris, A. A., & Wakemanxya, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional Children, 74*(4), 407-432.

Burns, M. K., Coddling, R. S., Boice, C. H., & 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*(1), 69.

Coddling, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice, 26*(1), 36-47.

Methe, S. A., Kilgus, S. P., Neiman, C., & Riley-Tillman, T. C. (2012). Meta-analysis of interventions for basic mathematics computation in single-case research. *Journal of Behavioral Education, 21*(3), 230-253.

Templeton, T. N., Neel, R. S., & Blood, E. (2008). Meta-analysis of math interventions for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 16*(4), 226-239.
<http://dx.doi.org/10.1177/1063426608321691>

[Reason for exclusion: All five meta-analyses above synthesise only studies involving single-case designs.]

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject design intervention research for students with LD. *Journal of Learning Disabilities, 33*, 114-136.

[Reason for exclusion: Most of the included studies were of language interventions (reading, writing, etc.) and the authors devote very little attention to mathematics specifically.]

Systematic reviews

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools (NCEE 2009-4060)*. Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Educational Sciences, US Department of Education.

Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning*. London: Nuffield Foundation.

Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J. (2010). *Developing effective fractions instruction for*

kindergarten through 8th grade: A practice guide (NCEE #2010-4039). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from whatworks.ed.gov/publications/practiceguides.

Other references

- Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296-233). New York: Macmillan.
- DfEE. (1999). *The National Numeracy Strategy: Framework for teaching mathematics from Reception to Year 6*. London: Department for Education and Employment.
- Fuson, K. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 243-295). New York: Macmillan.
- Hart, K. (Ed.) (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- Hodgen, J., & Marks, R. (2013). *The Employment Equation: Why our young people need more maths for today's jobs*. London: The Sutton Trust.
- Lamon, S. J. (2007). Rational Numbers and Proportional Reasoning: Toward a Theoretical Framework for Research. In F. K. J. Lester (Ed.), *Second handbook of Research on mathematics teaching and learning* (pp. 629-667). Greenwich, CT: Information Age Publishing.
- Ng, S. F., & Lee, K. (2009). The Model Method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, 40(3), 282-313.
- Ryan, J., & Williams, J. (2007). *Children's mathematics 4-15: learning from errors and misconceptions*. Buckingham: Open University Press.
- Sowder, J. (1992). Estimation and number sense. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371-389). New York: Macmillan.
- Thompson, I. (2001). British research on mental and written calculation methods for addition and subtraction. In M. Askew & M. Brown (Eds.), *Teaching and learning primary numeracy: Policy, practice and effectiveness. A review of British research for the British Educational Research Association in conjunction with the British Society for Research into Learning of Mathematics* (pp. 15-21). Southwell, Notts: British Educational Research Association (BERA).

8.4 Geometry

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of geometry and measures?

There are few studies that examine the effects of teaching interventions for and pedagogic approaches to the teaching of geometry. However, the research evidence suggests that representations and manipulatives play an important role in the learning of geometry. Teaching should focus on conceptual as well as procedural knowledge of measurement. Learners experience particular difficulties with area, and need to understand the multiplicative relations underlying area.

Strength of evidence: LOW

Findings

Geometry, measurement and spatial reasoning are important aspects of mathematics. In school geometry and measurement, students learn about the properties of points, lines, curves, surfaces and solids. Spatial reasoning is broader and includes things like the spatial orientation needed for everyday navigation as well as spatial visualisation, such as mental rotation.

Clements & Battista (1992) identified very few studies that examined the effect on attainment of teaching interventions and pedagogic approaches aimed at improving the learning of geometry and spatial reasoning (see also Battista, 1992). They did, however, highlight the important role of diagrams, representations and manipulatives in the learning of geometry. They also documented a number of key misconceptions (see also Dickson et al., 1984). For example, some children think that a square is not a square unless its base is horizontal. This suggests that teachers need to consider varying the orientation when presenting diagrams and examples to learners.

Clements & Battista (1992) highlight the promise of computers and technology to help develop geometric representations, but found little research investigating these effects. Battista's (2007) review, conducted 15 years later, documented a series of empirically-based theoretical studies that examined teaching and learning using LOGO and dynamic geometry software (DGS). Chan and Leung (2014) found a substantial positive effect ($d=1.02$) associated with the use of DGS, although more research is needed before assuming that DGS will be transformative in the classroom (Battista, 2007; Clements & Battista, 1992), particularly as the included studies were mostly small-scale and short-term (see also the Technology module).

Bryant's (2009) systematic review of the research on children's learning of geometry and spatial reasoning indicated that, whilst learners enter school with a great deal of implicit knowledge about spatial relations, they then have to learn how to represent this knowledge in language and symbols, which presents difficulties. The review recommended that teaching should focus on the conceptual basis of measurement, rather than just the procedural aspects, a finding also emphasised in Battista's (1992) review. This includes emphasising *transitive* relations (i.e., if $A < B$ and $B < C$, then $A < C$), and the idea of the iteration of standard units in measurement (e.g., tiling a rectangle with unit squares). Bryant (2009) makes clear links to the importance of the number line and the need to recognise that fractions and decimals expand the number system beyond whole numbers (see section on number). Learners encounter difficulties with area and need to understand the multiplicative relations underlying area. They will "understand this multiplicative reasoning better

when they first think of it as the number of tiles in a row times the number of rows than when they try to use a base times height formula” (p. 6) (see also Battista, 2007). Learners should also be encouraged to consider conservation (and equivalence) of area when adding, subtracting, and rearranging components of shapes to work out areas. Teachers should be aware that learners experience confusion when considering linear and area enlargements, and may incorrectly think that doubling the perimeter of a square or rectangle also doubles its area.

Evidence base

We found only one meta-analysis examining teaching interventions and pedagogic approaches relating to geometry, which addresses the effects of using DGS on attainment (Chan & Leung, 2014). However, the effect size may be inflated, because studies were largely small-scale and of short duration, and there may also have been novelty effects. As a result, for this section, we have also synthesised findings from three research reviews (Battista, 2007; Bryant, 2009; Clements & Battista, 1992).

Directness

We judge the evidence regarding children’s learning reported above to be relevant to England, although much of the work has been carried out in the US. However, since there are a very few relevant intervention studies, the findings are judged to have weak directness.

Threat to directness	Grade	Notes
Where and when the studies were carried out	1	Very few studies.
How the intervention was defined and operationalised	1	Very few studies.
Any reasons for possible ES inflation	1	Possible novelty factor; many studies are small-scale.
Any focus on particular topic areas	1	There is a pressing need for further research. Bryant (2009), for example, highlights a need for ‘basic’ research into various aspects of children’s learning of geometry and spatial relations.
Age of participants	1	Very few studies.

Overview of effects

Meta-analysis	Effect Size (<i>d</i>)	No of studies (<i>k</i>)	Quality	Comment
Chan & Leung (2014): Dynamic Geometry Software	1.02 [0.56, 1.48]	9	2	Short-term instruction with DGS significantly improved the achievement of primary learners $d = 1.82$ [1.38, 2.26], $k = 3$. The effect size may be inflated, because

[2002-2012]				studies were largely small-scale and of short duration.
-------------	--	--	--	---

References

Meta-analyses included

Chan, K. K., & Leung, S. W. (2014). Dynamic geometry software improves mathematical achievement: Systematic review and meta-analysis. *Journal of Educational Computing Research*, 51(3), 311-325.

Systematic reviews included

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. J. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 843-908). Greenwich, CT: Information Age Publishing.

Bryant, P. (2009). Paper 5: Understanding space and its representation in mathematics. In T. Nunes, P. Bryant, & A. Watson (Eds.), *Key understandings in mathematics learning*. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 4 December 2017.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 420-464). New York: Macmillan.

Systematic reviews excluded

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). *Teaching math to young children: A practice guide (NCEE 2014-4005)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. [Judged not to be relevant, because the guidance applies to younger children. One recommendation refers to geometry ("Teach geometry, patterns, measurement, and data analysis using a developmental progression"), but was judged to be at too general a level.]

8.5 Probability and Statistics

What is the evidence regarding the effectiveness of teaching approaches to improve learners' understanding of probability and statistics?

There are very few studies that examine the effects of teaching interventions for and pedagogic approaches to the teaching of probability and statistics. However, there is research evidence on the difficulties that learners experience and the common misconceptions that they encounter, as well as the ways in which they learn more generally. This evidence suggests some pedagogic principles for the teaching of statistics.

Strength of evidence: LOW

Findings

The reviews of research identified very few studies that examined the effect on attainment of teaching interventions and pedagogic approaches aimed at improving the learning of probability and statistics (Bryant & Nunes, 2012; Jones, Langrall & Monney, 2007; Shaughnessy, 1992, 2007). However, these research reviews do provide evidence on the difficulties that learners experience and the common misconceptions that they develop, as well as the ways in which they learn more generally.

Bryant & Nunes (2012) identify four cognitively demanding aspects to the learning of probability:

- Understanding randomness

- Working out the sample space

- Comparing and quantifying probabilities

- Understanding associations (and non-associations) between events

Drawing on his review of research, Shaughnessy (2007) outlines implications for teaching statistics:

- Build on students' intuitive notions of centre and variability

 - Emphasise variation and variability as key concepts in statistics (alongside the concept of central tendency)

 - Introduce comparison of data sets early in children's education, prior to the introduction of formal statistics

 - Help learners to understand the role of proportional reasoning in connecting populations and samples

- Highlight the importance of contextual issues in statistics

Although these implications are not strongly supported by evidence from intervention studies or teaching experiments, they nevertheless appear reasonable and are generally in line with pedagogic recommendations outlined elsewhere in this review.

Evidence base

We found no meta-analyses examining teaching interventions and pedagogic approaches relating to probability and statistics. As a result, for this section we have also considered findings from four research reviews (Bryant & Nunes, 2012; Jones, Langrall, & Monney, 2007; Shaughnessy, 1992, 2007).

Directness

We judge the evidence regarding children's learning reported above to be relevant to England, although much of the work has been carried out in the US. However, since there are a very few relevant intervention studies, the findings are judged to have weak directness.

Threat to directness	Grade	Notes
Where and when the studies were carried out	1	Very few studies.
How the intervention was defined and operationalised	1	Very few studies.
Any reasons for possible ES inflation	1	Very few studies.
Any focus on particular topic areas	1	There is a pressing need for further research.
Age of participants	1	Very few studies.

References

Meta-analyses included

None

Systematic reviews included

Bryant, P., & Nunes, T. (2012). Children's understanding of probability: A literature review. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 4 December 2017.

Jones, G. A., Langrall, C. W., & Monney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. J. Lester (Ed.), *Second handbook of Research on mathematics teaching and learning* (pp. 909-955). Greenwich, CT: Information Age Publishing.

Shaughnessy, J. M. (2007). Research on Statistics Learning and Reasoning. In F. K. J. Lester (Ed.), *Second handbook of Research on mathematics teaching and learning* (pp. 957-1010). Greenwich, CT: Information Age Publishing.

Shaughnessy, J. M. (1992). Research in probability and statistics: reflections and directions. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan.

9 Wider School-Level Strategies

9.1 Grouping by attainment or ‘ability’

What is the evidence regarding ‘ability grouping’¹⁶ on the teaching and learning of maths?

Setting or streaming students into different classes for mathematics based on their prior attainment appears to have an overall neutral or slightly negative effect on their future attainment, although higher attainers may benefit slightly. The evidence suggests no difference for mathematics in comparison to other subjects. The use of within-class grouping at primary may have a positive effect, particularly for mathematics, but if used then setting needs to be flexible, with regular opportunities for group reassignment.

Strength of evidence: MEDIUM

Findings

Grouping by ‘ability’ is a common organisational structure in both primary and secondary schooling. It may take a number of forms, sometimes used in combination (definitions taken from Marks, 2016, p. 4):

Setting: children are placed into ability classes for particular subjects (e.g., all Year 8 pupils are grouped into different classes for mathematics); a child could be in different sets for different subjects.

Streaming: children are placed in the same ability classes for all subjects based on general ability. This is often referred to as ‘tracking’ in the US.

Within-class grouping: children are allocated to table groups within the class for all or some subjects, based on general ability or subject-specific ability.

Mixed-ability: classes are not grouped by ability and in a multi-form entry school each class in a year-group should contain the same range of attainment.

In the US, there are also specific grouping programmes involving cross-grade / vertical subject grouping. This is uncommon in England.

There is a large research base concerning ability grouping and it continues to be a ‘hot topic’ in mathematics education. This may be due to concerns over managing the wide range of attainment within year groups, although Brown et al. (1998, pp. 371-2), in reviewing evidence related to the instigation of the National Numeracy Strategy, note that “countries that have the largest standard deviations are exactly those of the Pacific rim, like Japan and Korea, which teach unsorted classes on an undifferentiated curriculum.”

The literature base for ability grouping not only includes a number of primary studies but also an unusually large number of meta-analyses and research syntheses.

These have now been further synthesised by two 2nd-order meta-analyses (syntheses of the meta-analyses). For the purpose of this module, we focus on these two 2nd-order analyses (Steenbergen-Hu et al., 2016; EEF, 2017), which bring

¹⁶We maintain the nomenclature of the majority of the literature, using the term ‘ability-grouping’, although we recognise the contested nature of this term.

together 15 meta-analyses based on 172 primary studies (see evidence base below).

The pooled effect for between-class grouping (setting and streaming) suggests an overall neutral or slightly negative effect on attainment. However, higher attainers may gain slightly from the practice.

The evidence base for within-class grouping (usually seen in primary schools) is limited, but suggests positive effects for mathematics. Lou et al. (1996) found that the effects of within-class grouping for mathematics and science combined ($d=0.20$) and for reading and liberal arts ($d=0.13$) were significantly greater than for other subjects. These positive effects should be treated with a degree of caution, however; Slavin (1987) suggests that the positive effect may be a feature of the flexibility of such classroom organisation structures, which may allow learners to frequently move between groups in response to their changing needs, even though, in practice, such movement may be limited.

Differentiated grouping may widen the attainment spread. The picture is more complicated, moderated by grouping type, attainment level, flexibility and subject, as can be seen in the EEF Toolkit and Steenbergen-Hu et al. (2016) discussions.

Evidence base

We base this module on two 2nd order analyses: Steenbergen-Hu et al. (2016) and the EEF Toolkit strand: setting or streaming.

Steenbergen-Hu et al. (2016) draw on 13 meta-analyses in their second-order meta-analysis, 11 of which reviewed the academic effects of between-class grouping. This analysis is deemed to be of high methodological quality, but is based on a synthesis of 13 meta-analyses that Steenbergen-Hu et al. judge to be either of medium or low quality. These meta-analyses are themselves based on the syntheses of studies, some of which contained methodological and reporting flaws (and, in particular, very few studies involved random assignment). Of these 13 meta-analyses, it was clear in only three (Slavin, 1987, 1993; Lou et al., 1996) that a subject-moderator analysis examining the specific impact of ability grouping in mathematics had been conducted. Slavin (1993) reports no differences for mathematics at the middle school level, while Slavin's 1987 study suggests results that are inconclusive for setting just for mathematics in primary. Lou et al.'s (1996) study combined mathematics with science and did not involve between-class grouping. It should be noted that it was not possible to determine how the moderator analysis had been conducted for Slavin's studies.

The 13 meta-analyses drew on 643 primary-studies, of which the authors found 172 to be unique. Of these, we estimate that 20% (i.e. approximately 35 studies) are specifically related to mathematics, while mathematics is likely to form an element of the general studies, which form approximately 60% of this literature, although it is not possible to disaggregate the effects on different subjects for many of these studies.

The EEF 'Setting or Streaming' toolkit draws on six meta-analyses (in addition to a range of single studies and reviews). Four of these also appear in Steenbergen-Hu et al. (2016). The two not included are less applicable to our review: Gutierrez and Slavin (1992) examine cross-grade programmes, while Puzio and Colby's (2010) study examines reading and within-class grouping.

The effect sizes found in the 15 meta-analyses are shown in the table below. This is based on the data extracted by Steenbergen-Hu et al. (2016) for their 13 included meta-analyses and from the original papers for Gutierrez and Slavin (1992) and Puzio and Colby (2010). It should be noted that for the four common meta-analyses the reported effect sizes do not always correspond; this may be due to reporting for a particular sub-group.

Meta-analysis	Steen- berg- en-Hu et al. (2016)	EEF Too lkit	ES	k	Comments
Goldring, E. B. (1990). Assessing the status of information on classroom organizational frameworks of gifted students. <i>Journal of Educational Research</i> , 83, 313–326. doi:10.1080/00220671.1990.10885977	✓		0.35	18	ES for gifted students overall
Henderson, N. D. (1989). A meta-analysis of ability grouping achievement and attitude in the elementary grades (Unpublished doctoral dissertation). Mississippi State University, MS.	✓		-0.30	4	Overall ES
			0.02	2	High ability
			-0.00 014	2	Medium ability
Kulik, C. C. (1985, August). Effects of inter-class ability grouping on achievement and self-esteem. Paper presented at the 93rd annual convention of the American Psychological Association, Los Angeles, CA.	✓		0.09	78	Overall ES
			0.12	≤7 4	Medium- ability
			0.12	4	Low ability
Kulik, C. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. <i>American Educational Research Journal</i> , 19, 415–428. doi:10.3102/00028312019003415	✓	✓	0.10	51	Overall ES
			0.02	33	Medium- ability
			0.02	4	Low ability
Kulik, C. C., & Kulik, J. A. (1984). Effects of ability	✓	✓	0.19	28	Overall ES

Meta-analysis	Steen- berg- en-Hu et al. (2016)	EEF Too lkit	ES	k	Comments
grouping on elementary school pupils: A meta-analysis. Paper presented at the annual meeting of the American Psychological Association, Toronto, Ontario, Canada.			0.02	19	Medium-ability
Kulik, J. A., & Kulik, C. C. (1987). Effects of ability grouping on student achievement. <i>Equity & Excellence in Education</i> , 23(1–2), 22–30. doi:10.1080/1066568870230105	✓		0.06	49	Overall ES
			0.12	40	High ability
			0.04	33	Medium ability
			0.00	39	Low ability
Kulik, J. A., & Kulik, C. C. (1992). Meta-analytic findings on grouping programs. <i>Gifted Child Quarterly</i> , 36, 73–77. doi:10.1177/001698629203600204	✓		0.03	51	Overall ES
			0.10	36	High ability
			-0.02	36	Medium ability
			-0.01	36	Low ability
Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. <i>Review of Educational Research</i> , 66, 423–458. doi:10.3102/00346543066004423	✓	✓	0.17	51	Overall ES (N.B. for within-class grouping)
			0.27	≤18	High ability (N.B. for within-class grouping)
			0.18	≤11	Medium ability (N.B. for within-class grouping)
			0.36	≤24	Low ability (N.B. for within-class grouping)
Mosteller, F., Light, R. J., & Sachs, J. A. (1996). Sustained	✓		0.00	10	Overall ES
			0.08	10	High ability

Meta-analysis	Steen- berg- en-Hu et al. (2016)	EEF Too lkit	ES	k	Comments
inquiry in education: Lessons from skill grouping and class size. <i>Harvard Educational Review</i> , 66, 797– 842. doi:10.17763/haer.66.4.36m328762x21610x			-0.04	10	Medium ability
			-0.06	10	Low ability
Noland, T. K. (1986). The effects of ability grouping: A meta-analysis of research findings. Retrieved from http://eric.ed.gov/?id=ED269451	✓		0.01	≤5 0	Overall ES
			0.16	≤5 0	High ability
			-0.45	≤5 0	Medium ability
			0.18	≤5 0	Low ability
Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. <i>Review of Educational Research</i> , 57, 293–336. doi:10.3102/00346543057003293	✓		-0.54	14	
Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. <i>Review of Educational Research</i> , 60, 471–499. doi:10.3102/00346543060003471	✓	✓	-0.03	29	Overall ES
			-0.02	15	High ability
			-0.07	15	Medium ability
			-0.03	15	Low ability
Slavin, R. E. (1993). Ability grouping in the middle grades: Achievement effects and alternatives. <i>Elementary School Journal</i> , 93, 535–552. doi:10.1086/461739	✓		-0.01	27	Overall ES
			0.01	14	High ability
			-0.07	14	Medium ability
			-0.02	14	Low ability
Gutierrez, R., & Slavin, R. E. (1992). Achievement Effects of the Non-graded Elementary		✓	0.46	9	Joplin like non-graded programs.

Meta-analysis	Steen- berg- en-Hu et al. (2016)	EEF Too lkit	ES	<i>k</i>	Comments
School: A Retrospective Review.			0.34	14	Non-graded Programs Involving Multiple Subjects (Comprehensive Programs)
Puzio, K., & Colby, G. (2010). The Effects of within Class Grouping on Reading Achievement: A Meta-Analytic Synthesis. Society for Research on Educational Effectiveness.		✓	0.22	15	<small>Within-class</small> grouping interventions

Directness

The 15 meta-analyses were published between 1982 and 2010. Seven were published in the 1980s and seven in the 1990s. This suggests that the literature may be somewhat dated-

The majority of the literature is based in the US. Although ability grouping systems do differ and have different labels, we judge that there are still enough similarities for this literature to be applicable to the context of England. Single studies in England tend to confirm the applicability of the results from the US literature.

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Most studies were carried out in the US; however studies in England tend to confirm the findings.
How the intervention was defined and operationalised	2	Some differences in terms used.
Any reasons for possible ES inflation	2	
Any focus on particular topic areas	3	
Age of participants	3	

Further research

The largely neutral effects of ability grouping are surprising to many teachers and other professionals in education, and this is particularly so for mathematics. Given this, and the widespread use of ability grouping in school mathematics, it is important to better understand how teachers and schools should best group students so as to

address the needs of students at all attainment levels in mathematics. There is scope for further analysis and research, both in terms of impact and alternatives. In particular, we judge that there is a need to investigate the effects of different combinations of approaches to addressing the different needs of students at different levels of attainment. This is of particular importance in the light of the evidence on cooperative learning (see module). As Slavin (1993) observed, “Revisiting individualized instruction or mastery learning in the context of untracking middle schools may be fruitful ... combining individualization with cooperative learning has turned out to be an effective strategy in mathematics in the upper-elementary grades and is likely to be useful in the middle grades as well” (p. 547). There is also a need to better understand within-class grouping at the primary level, in addition to developing our understanding of the impacts of all forms of ability-grouping on equity in the teaching and learning of mathematics.

Overview of 2nd-order meta-analysis reported effects

2 nd -order meta-analysis	Variable	Effect Size (d)
Steenbergen-Hu et al. (2016)	Within-class grouping	$0.19 \leq g \leq 0.30$
	Cross-grade subject grouping	0.26
	Between class grouping	$0.04 \leq g \leq 0.06$
EEF Toolkit	Setting and streaming (low-attainers)	-0.09

References

2nd Order Meta-analyses included

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K–12 Students’ Academic Achievement: Findings of Two Second-Order Meta-Analyses. *Review of Educational Research*, 86(4), 849-899.

Education Endowment Foundation (2017) *Teaching & Learning Toolkit: Setting or streaming*. London: EEF.

Other references

Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385

Higgins, S., Katsipatakis, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., ... & Mason, D. (2016). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*. London: Education Endowment Foundation.

- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66, 423–458.
- Marks, R. (2016). *Ability grouping in primary schools: case studies and critical debates*. Northwich: Critical Publishing
- Slavin, R. E. (1993). Ability grouping in the middle grades: Achievement effects and alternatives. *The Elementary School Journal*, 93(5), 535-552.
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57, 293–336.

9.2 Homework

What is the evidence regarding the effective use of homework in the teaching and learning of mathematics?

The effect of homework appears to be low at the primary level and stronger at the secondary level, although the evidence base is weak. It seems to matter more that homework encourages students to actively engage in learning rather than simply learning by rote or finishing off classwork. In addition, the student's effort appears to be more important than the time spent or the quantity of work done. This would suggest that the teacher should aim to set homework that students find engaging and that encourages metacognitive activity. For primary students, homework seems not to be associated with improvements in attainment, but there could be other reasons for setting homework in primary, such as developing study skills or student engagement. Homework is more important for attainment as students get older. As with almost any intervention, teachers make a huge difference. It is likely that student effort will increase if teachers value students' homework and discuss it in class. However, it is not clear that spending an excessive amount of time marking homework is an effective use of teacher time.

Strength of evidence: LOW

Findings

Homework involves a variety of tasks assigned by teachers for pupils to complete – usually independently – outside of school hours (Pattall et al., 2008). At the primary level, this often involves reading, and practising spellings and number facts, such as multiplication tables (Higgins et al., 2013). At secondary level, homework often includes preparation for upcoming lessons, completing work not finished in lessons, revision activities and extended projects.

While it has been suggested that increasing the quantity of or challenge associated with homework may plausibly be a strategy for raising standards in primary mathematics (e.g., Brown et al., 1998), the current evidence – as outlined in the EEF Toolkit – suggests that the effect of homework on general academic achievement is low at the primary level and stronger, but with wide variation, at the secondary level (Higgins et al., 2013). However, the evidence is weak and not entirely consistent (see evidence base below). On the basis of six experimental studies – of which only one was in mathematics – Cooper et al. (2006) report an ES of 0.60. Paschal et al.'s (1984) synthesis of a set of older experimental studies found higher effects for homework amongst primary students (Year 5 and Year 6) compared to secondary. On the other hand, Cooper et al.'s (2006) meta-analysis of correlational studies found *no* effect for primary ($r = -0.04$), compared to a medium-sized effect for secondary ($r = 0.25$). This concurs with the findings of the Canadian Council on Learning's (CCL, 2009) systematic review of the impact of homework on academic achievement, which again was not focused on mathematics. Based on 10 recent studies, the review found evidence that the use of homework increases achievement to a moderate degree (particularly with older pupils and lower-attainers). However, the evidence is varied and contains some contradictory findings. They argue that homework is a diverse activity, which has the potential to impact positively, or negatively, on attainment. Their findings also suggest that homework which

promotes 'active learning' (such as metacognition) rather than "rote repetition of classroom material" is more likely to increase attainment (p.44).

Looking at mathematics specifically, the evidence is somewhat contradictory. In an analysis of a longitudinal US dataset, based on a cohort of approximately 25,000 students in Grade 8 (Year 9) in 1988, Eren and Henderson (2011, p.960) found that mathematics was the only subject with a "consistently and statistically meaningful large effect on test scores", although both Paschal et al.'s (1984) and Cooper et al.'s meta-analyses found no significant differences between different school subjects.

The Canadian Council on Learning (CCL, 2009) found that the quality of the homework task and the level of student engagement seemed to be more important than the amount of time a student spent on homework. For example, Trautwein (2007) reported on three studies with Grade 8 (Year 9) students in Germany, based on an analysis of the TIMSS 1995 and PISA 2000 data, and an associated longitudinal study. This analysis suggested that, for mathematics, effort put into homework, rather than the amount of time spent on it, was associated with attainment gains.

There is also something to be understood about the effects of technology-based homework in mathematics, with Steenbergen-Hu and Cooper's (2013) finding that intelligent tutoring systems (ITS) appear to be more effective than pencil-and-paper homework assignments in mathematics, although this was based on very limited evidence, and the overall effect of ITS was small. Interestingly, Eren and Henderson (2011, p. 960) found that "the teachers' treatment of the homework (whether it is being recorded and/or graded) does not appear to affect the returns to math homework", although there is obvious caution to be advised in how this single-study finding is implemented by practitioners.

The evidence of the efficacy of after-school programmes is slightly stronger, and Crawford's (2011) meta-analysis reported an ES of $d=0.42$ for mathematics, based on a synthesis of 10 studies. However, moderator analysis suggested that any impact would be dependent on the design of the after-school programme.

Evidence base

We found very limited evidence regarding the use of homework in mathematics specifically. Given the limited evidence base, we have drawn on syntheses of correlational studies, together with some recent single studies, in order to supplement the meta-analyses and systematic reviews. We would advise caution in interpreting and applying findings drawn from these studies.

We have included two meta-analyses considering the effect of homework on attainment, although these consider attainment in general rather than mathematics specifically. These meta-analyses contain only a small number of experimental studies in mathematics, and few of these are either robustly designed or have been conducted recently. Societal changes outside school are of particular relevance to homework, because it is possible that young people are less or more willing to engage in homework in the present day than they were 40 or 50 years ago.

The findings draw heavily on correlational studies, which provide evidence of associations but not of causation. Hence, any positive effects associated with homework may be the result of other factors.

There is certainly a need for future research specifically examining the case of mathematics across both primary and secondary aged-pupils, providing guidance on the most effective uses of homework in mathematics and identifying the causal relationships between homework and mathematical attainment. However, as Cooper et al. (2006) indicate, such research will need to draw on a variety of research designs and methodologies, partly because of inherent difficulties in conducting robust experimental studies involving homework, including the difficulty of withholding from some students any intervention, such as homework, which is widely presumed to have benefits.

Directness

Within the limited evidence, it seems clear that homework is poorly understood and therefore detailed guidance is limited, although secondary students and low attainers seem likely to gain more. However, as the Canadian Council on Learning (CCL, 2009) conclude, the evidence suggests that useful principles for teachers are to design homework that requires, or encourages, students to engage in active learning (rather than simple repetition of classroom material). Since student effort is more important than time spent on homework, it would seem beneficial to value effort and to set tasks that are likely to engage all students more.

Threat to directness	Grade	Notes
Where and when the studies were carried out	1	The studies drawn on in the meta-analyses are now fairly dated, and the educational / policy /societal context has changed. The vast majority of the studies were located in the US.
Strengths and weaknesses in the research design	1	Few of the studies had robust research designs. Cooper et al. (2006) highlight the inherent difficulties in conducting experimental studies involving homework, which make identifying causal relationships difficult.
How the intervention was defined and operationalised	2	Homework as an intervention is poorly defined.
Any reasons for possible ES inflation	2	Studies suggest a stronger effect for lower-attainers, but this is not accounted for in all primary studies (and may be inflated by the restricted attainment ranges in the samples). A further source of bias may be that homework interventions may be affected by confounding factors, such as compliance and other student behaviours.
Any focus on particular topic areas	2	Relatively few intervention studies are focused on mathematics.
Age of participants	2	Limited research at the primary level.

Overview of effects

Meta-analysis	Effect Size	No of studies (<i>k</i>)	Quality	Comment
<i>Effect of homework interventions on attainment</i>				
Cooper et al. (2006): attainment in general	$d = 0.60$ [0.38, 0.82]	6	3	Random effects model estimate for all 6 studies is reported (i.e. all experimental designs combined). However, only one of these 6 studies is in mathematics (Y3). The authors indicate that a great deal of caution should be exercised in interpreting this estimate, due to limitations in the number and robustness of the studies synthesised.
Paschal et al. (1984): mathematics attainment	$d = 0.23$	60 ESs (based on <15 reports)	1	The majority of effects considered were for mathematics (60 effects for mathematics out of a total of 81 effects, taken from 15 reports). The studies are now dated (1964-1980) and almost wholly conducted in the US. Studies were based on experimental designs. The 81 effects synthesised include 9 attitudinal effects. It is not clear whether Paschal et al. have taken dependencies between effects into account.
Paschal et al. (1984): attainment in general	$d = 0.36$	81 (based on 15 reports)	1	The synthesis finds significantly higher effects for Y5 and Y6.

<i>Effects based on correlations between homework and attainment</i>				
Cooper et al. (2006): primary	$r = -0.04$	10	3	
Cooper et al. (2006): secondary	$r = 0.25$	23	3	
<i>Effect of homework interventions using intelligent tutoring systems (ITS) compared to pencil and paper based homework</i>				
Steenbergen-Hu and Cooper (2013)	$g = 0.6$	2	3	Two small studies are cited, both in primary, with effects in favour of ITS of $g=0.61$ and 0.61 .
<i>Effect of after-school programmes on attainment</i>				
Crawford (2011)	$d=0.42$	10	2	ES for reading similar to mathematics ($d=0.38$).

References

Meta-analyses included

- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of educational research*, 76(1), 1-62. [k=50]
- Crawford, S. T. (2011). *Meta-Analysis of the Impact of After-School Programs on Students Reading and Mathematics Performance*. (ProQuest UMI 3486475 EdD), University Of North Texas.
- Paschal, R. A., Weinstein, T., & Walberg, H. J. W. (1984). The effects of homework on learning: A quantitative synthesis. *The Journal of Educational Research*, 78(2), 97-104. [k=15]
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970-987. doi:10.1037/a0032447

Other references

- Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385.
- Higgins, S., Katsipatakis, M., Kokotsaki, D., Coleman, R., Major, L. E., & Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*. London: Education Endowment Foundation.
- Canadian Council on Learning (2009). A systematic review of literature examining the impact of homework on academic achievement. Toronto: Canadian Council on Learning.

9.3 Parental engagement

What is the evidence regarding parental engagement and learning mathematics?

The well-established association between parental involvement and a child's academic success does not appear to apply to mathematics, and there is limited evidence on how parental involvement in mathematics might be made more effective. Interventions aimed at improving parental involvement in homework do not appear to raise attainment in mathematics, and may have a negative effect in secondary. However, there may be other reasons for encouraging parental involvement. Correlational studies suggest that parental involvement aimed at increasing academic socialization, or helping students see the value of education, may have a positive impact on achievement at secondary.

Strength of evidence: LOW

Findings

The EEF (2017) toolkit states that “The association between parental involvement and a child's academic success is well established” (EEF, 2017). However, Patall et al.'s (2008) meta-analysis of correlational evidence suggests that this association does not appear to hold for mathematics. They found a significant negative association between parental involvement and achievement in mathematics ($d = -.19$), compared to a significant positive association for reading ($d = .20$). They also found that association between parental involvement in homework and attainment was strong and positive for elementary-age pupils ($d = .22$) and strong and negative for middle-school students ($d = -.18$).

Patall et al. (2008) examined experimental studies looking at the impact on attainment of training parents to be involved in homework. Their findings are limited by the small number of studies (14, with 10 in mathematics), only some of which involved randomisation (9) or pre-tests (5). Their findings were mixed, with effects on attainment ranging from $d = .00$ to $d = .22$. Moderator analysis indicated that the effects were positive for elementary students and negative for middle school students, with no differences between mathematics and reading. Essentially, “the effect of training parents for homework involvement has at best a slightly positive overall impact on achievement” (Patall et al., 2008, p.1062).

Pattall et al. suggest that the negative effects for middle school may be due to many parents lacking the skills, knowledge and confidence needed to provide subject-specific support. Indeed, evidence from Brooks et al.'s (2008) systematic review suggests that improving parents' skills, knowledge and confidence is challenging, particularly in mathematics.

Hill & Tyson's synthesis of correlational effects found stronger association between general parental involvement and achievement in middle school, although the association was stronger for academic socialisation, or communicating the value of education ($r = 0.39$), than for home-based engagement, such as assisting with homework ($r = 0.03$).

Evidence base

As noted previously, while we identified two fairly recent meta-analyses, we draw predominantly on Patall et al. (2008), due to intervention overlap and methodological

quality. Patall et al. (2008) draw on 45 studies covering the period 1987-2004, although these are split across three separate analyses. While the overall effect sizes were small, there was quite substantial variation in effects across the studies, suggesting that some caution should be applied.

Directness

Threat to directness	Grade	Notes
Where and when the study was carried out	2	The studies in the meta-analysis were conducted in the US and Canada. The correlational effects for England are likely to be similar, so we do not regard this as a threat to directness. The interventions in the experimental studies in Patall et al. may not directly transfer, due to differences in societal factors between the US and England.
How the intervention was defined and operationalised	2	Parental engagement is clearly defined, although the parental engagement interventions are less clearly defined.
Any reasons for possible ES inflation	2	Publication bias may mean these ESs are over-estimates. Many of the studies are not robust and few have a pre-test. Most of the studies are correlational, so provide evidence of associations not causation.
Any focus on particular topic areas	3	
Age of participants	3	Grades 1 – 8

Overview of effects

Meta-analysis	Effect Size	No of studies (k)	Quality	Comment
<i>Overall effect of parent training for homework involvement on outcomes</i>				
Patall et al. (2008)	0.00 [-0.27, 0.27]	3	3	Unadjusted ES from random effects model of randomized experiments without pre-tests. This analysis excluded two studies as outliers. Larger, but n.s., positive effects

				were found for all 5 studies (0.09, 95%CI [-0.16, 0.34])
Patall et al. (2008)	0.22 [0.01, 0.43]	5	3	Adjusted (to include pre-test data) ES from random effects model of quasi-experiments
<i>Moderator analyses examining the effect of parent training for homework involvement on academic achievement by grade level</i>				
Patall et al. (2008); Elementary	0.23 [-0.06, 0.52]	3	3	Random effects model
Patall et al. (2008); Middle school	-0.18 [-0.49, 0.14]	2	3	Random effects model
<i>Moderator analyses examining the effect of parent training for homework involvement on academic achievement by subject area</i>				
Patall et al. (2008); Mathematics	0.12 [-0.23, 0.47]	4	3	Random effects model
Patall et al. (2008); Reading	0.09 [-0.26, 0.44]	2	3	Random effects model
<i>Correlational evidence on the association between parental involvement and attainment</i>				
Patall et al. (2008); Mathematics	-.19 [-.24, -.15]	3	3	Random effects model
Patall et al. (2008); Reading	.13 [-.25, .48]	6	3	Random effects model
Patall et al. (2008); Language Arts	.12 [.05, .20]	3	3	Random effects model
Hill & Tyson (2009); academic socialisation	.39 [.26, .44]	16	3	Random effects model
Hill & Tyson (2009); help with homework	-.11 [-.25, -.04]	6	3	Random effects model
Hill & Tyson (2009); activities at home	.12 [.05, .19]	5	3	Random effects model

References

Meta-analyses included

- Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent Involvement in Homework: A Research Synthesis. *Review of Educational Research*, 78(4), 1039-1101.
- Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: a meta-analytic assessment of the strategies that promote achievement. *Developmental psychology*, 45(3), 740. [We have not considered the analysis of interventions aimed at improving parental involvement since all 5 interventions considered were included in Patall's synthesis.]

Other references

- Anthony, G., & Walshaw, M. (2007). *Best evidence synthesis: Effective pedagogy in Pangarau/Mathematics*. Wellington, NZ: Ministry of Education. (#128)
- Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, 103(1), 51-73. (#25)
- Brooks, G., Pahl, K., Pollard, A., & Rees, F. (2008). Effective and inclusive practices in family literacy, language and numeracy: a review of programmes and practice in the UK and internationally. Reading: CfBT.
- Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385 (#127)
- Canadian Council on Learning. (2009). A systematic review of literature examining the impact of homework on academic achievement. Toronto: Canadian Council on Learning.
- Cara, O., & Brooks, G. (2012). Evidence of the Wider Benefits of Family Learning: A Scoping Review. BIS Research Paper Number 93. London: Department for Business, Innovation and Skills.
- Education Endowment Foundation (2017) *Teaching & Learning Toolkit: Parental engagement*. London: EEF.
- OECD. (2004). *Program for International Student Assessment. Learning for tomorrow's world: First results from PISA 2003* (OECD Publications No. 53799 2004). Paris: Organisation for Economic Cooperation and Development.

10 Attitudes and Dispositions

How can learners' attitudes and dispositions towards mathematics be improved and maths anxiety reduced?

Positive attitudes and dispositions are important to the successful learning of mathematics. However, many learners are not confident in mathematics. There is limited evidence on the efficacy of approaches that might improve learners' attitudes to mathematics or prevent or reduce the more severe problems of maths anxiety. Encouraging a growth mindset rather than a fixed mindset is unlikely to have a negative impact on learning and may have a small positive impact.

Strength of evidence: LOW

Findings

In Section 3 of this document, we described how attitudes and dispositions are important to learning and doing mathematics. In a meta-analysis of US studies, Ma and Kishnor (1997) found that attitudes appear to have a small causal effect on attainment ($r=0.08$), whereas the opposite appears not to be the case. However, the meta-analysis was based on causal modelling of just five, albeit large, naturalistic studies. Ma & Kishnor's (1997) finding suggests that improving student attitudes towards mathematics may have a small impact on attainment.

International survey evidence appears to contradict the common view that attitudes are more negative in England in comparison to other countries internationally. Evidence from the latest TIMSS and PISA surveys indicate that attitudes to mathematics amongst learners in England are above the international average and similar to those of the highest-attaining countries. Attitudes follow the general international pattern in declining over time (see Section 3). In TIMSS 2015, the overall proportion of learners who were either confident or very confident in mathematics was 80% at Year 5 and 65% at Year 9 (Greany et al., 2016). In PISA 2012, at age 15, almost all learners in England agreed or strongly agreed with the statement, "If I put in enough effort I can succeed in mathematics" (96% compared to an international average of 92%) (Wheater et al., 2014). However, the international studies indicate that, amongst learners within England and other countries, there is a relationship between attitudes and attainment, with lower attainers tending to have more negative attitudes. The TIMSS 2015 survey collected evidence on confidence and enjoyment (or liking mathematics) as well as whether learners valued mathematics or perceived their mathematics teaching to be engaging. This evidence indicates that, in England and internationally, the association between student attainment and attitudes was strongest for confidence and enjoyment, particularly at Year 9 (Greany et al., 2016).

We found surprisingly little evidence demonstrating effective approaches to improving attitudes. Muenks and Miele's (Forthcoming) research synthesis examined learners' perceptions of the relationship between effort and ability. They found that some teacher actions, such as a challenge to "think deeply" (Middleton & Midgeley, 2002, p. 386) and the promotion of an incremental, or malleable, theory of intelligence, appear to encourage learners to believe that increased effort will increase their own abilities, whereas social comparison and competition tend not to encourage a positive relationship (see also Middleton & Spanias, 1999). Lazowski and Hulleman's (2016) meta-analysis found a moderate ES for a range of research-

based approaches aimed at increasing motivation ($d=0.49$), although these effects were across school subjects rather than mathematics-specific. They concluded by suggesting that the benefits of motivational interventions may potentially be considerable at minimal cost. However, they observe that existing approaches have largely only been evaluated in experimental settings and that translating these experimental approaches into research-based interventions that can be implemented by teachers is at a very early stage of development. (See also Metacognition and Parental Engagement modules for related strategies.)

In recent years, the importance of learners adopting a growth mindset has been widely promoted by teachers and schools in England, particularly in mathematics (Boaler 2013; see Simms, 2016, for a critique). Muenks and Miele (Forthcoming) suggest that some growth mindset interventions appear promising. However, this intervention-based research is at a very early stage of development, and, whilst some studies have shown small benefits for some learners (e.g., Paunesku et al., 2015), other studies have not shown statistically significant benefits (e.g., Churches, 2016; Paunesku et al., 2011a, 2011b; Rienzo et al., 2015). This suggests that the promotion of a growth mindset is unlikely to have a negative impact on learning and may have a small positive impact in some contexts.

Maths anxiety is defined as “a feeling of tension and anxiety that interferes with the manipulation of numbers and the solving of mathematical problems” (Richardson & Suinn, cited in Dowker et al., 2016, p. 1). Although correlated with attitudes and learner self-concept, maths anxiety is distinct from attitudes, such as confidence in and liking of mathematics. Maths anxiety has a larger detrimental impact on attainment than attitudes in general, by disrupting working memory and through avoidance of mathematical activities (Dowker et al., 2016; see also studies cited in Dowker et al., 2016, including Ma, 1999). However, in their synthesis of the research evidence, Dowker et al. conclude that the causal relationships between maths anxiety and attainment are not well understood and, whilst there is some promising research, there is only a limited understanding of how to reduce maths anxiety. Hembree’s (1990) meta-analysis indicated some promising approaches to reducing maths anxiety and raising attainment, including systematic desensitisation, or graduated exposure therapy. However, these approaches were largely evaluated with college students in the US and do not provide practical guidance for mathematics classrooms in England. Whilst Dowker et al. (2016) highlight some promising approaches to addressing maths anxiety, these are at an early stage of development and more research is needed to address this issue.

Evidence base

As noted in the findings, the evidence base on approaches either to improving attitudes and dispositions or to reduce maths anxiety is very limited.

Meta-analysis	Focus	k	Quality	Date Range
Hembree (1990)	Maths anxiety	13	2	Not given
Lazowski & Hulleman (2016)	Motivation	92	3	Prior to May 2015

Ma & Kishor (1997)	Relationship between attitude and attainment	113	2	1966-1993
--------------------	--	-----	---	-----------

Directness

Threat to directness	Grade	Notes
Where and when the studies were carried out	2	Many of the studies were carried out in the US.
How the intervention was defined and operationalised	1	All the meta-analyses and syntheses comment that, whilst studies are promising, much more work needs to be done to enable implementation with fidelity by teachers.
Any reasons for possible ES inflation	2	Many of the interventions were delivered by researchers rather than in regular classrooms.
Any focus on particular topic areas	2	
Age of participants	2	Many of the original studies in Hembree (1990) were carried out with college students.

Overview of effects

Meta-analysis	Effect Size (d)	No of studies (k)	Comment
<i>Effect of interventions to increase motivation on attainment</i>			
Lazowski & Hulleman (2016)	0.49 95% CI [0.43, 0.56]	92	Synthesises a range of interventions based on different theoretical approaches, all aimed at improving motivation. (However, this is across subjects in general; i.e., not focused on mathematics, and there is no moderator analysis for different subjects). Average effects for different approaches varying from d=0.36 to d=0.74.
<i>Effect of interventions to reduce maths anxiety on attainment and maths anxiety</i>			
Hembree (1990)	See comments.	-	<i>Effects on maths anxiety:</i> d=-1.04 (Systematic desensitisation only, k=18) d=-0.51 (Cognitive restructuring only, k=14) d=-1.15 (Cognitive-behavioural approaches, ie both together, k=10)

			<i>Effects on attainment:</i> $d=0.60$ (Systematic desensitisation only, $k=12$) $d=0.32$ (Cognitive restructuring only, $k=7$) $d=0.50$ (Cognitive-behavioural approaches, i.e. both together, $k=4$)
<i>Relationship between attitude and attainment in mathematics</i>			
Ma & Kishor (1997)	See comments.		<i>Correlation between attitudes and attainment:</i> $r=0.12$, 95% CI [0.12, 0.13], $k=107$ ($N=59,925$). <i>Causal relationship attitudes to attainment:</i> $r=0.08$, 95% CI [0.07, 0.09], $k=5$ ($N=20,227$). <i>Causal relationship attainment to attitudes:</i> $r=0.00$, 95% CI [-0.01, 0.01], $k=5$ ($N=20,227$).

References

Meta-analyses included

- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(1), 33-46.
- Lazowski, R. A., & Hulleman, C. S. (2016). Motivation Interventions in Education. *Review of Educational Research*, 86(2), 602-640. doi:doi:10.3102/0034654315617832
- Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: a meta-analysis. *Journal for Research In Mathematics Education*, 28(1), 26-47.

Secondary meta-analysis

- Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, 30(5), 520-540.

Research syntheses

- Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics Anxiety: What Have We Learned in 60 Years? *Frontiers in Psychology*, 7(508). doi:10.3389/fpsyg.2016.00508.
- Middleton, J. A., & Spanias, P. A. (1999). Motivation for achievement in mathematics: Findings, generalizations and criticisms of the research. *Journal for Research in Mathematics Education*, 30(1), 65-88.
- Muenks, K., & Miele, D. B. (Forthcoming, online first). Students' Thinking About Effort and Ability. *Review of Educational Research*, 0034654316689328. doi: 10.3102/0034654316689328

Other references

- Boaler, J. (2013). Ability and Mathematics: the mindset revolution that is reshaping education. *Forum*, 55(1), 143-152.
- Churches, R. (2016). *Closing the gap: test and learn*. London: Department for Education.
- Greany, T., Barnes, I., Mostafa, T., Pesniero, N., & Swenson, C. (2016). *Trends in Maths and Science Study (TIMSS): National Report for England*. London: Department for Education.
- Middleton, M. J., & Midgley, C. (2002). Beyond Motivation: Middle School Students' Perceptions of Press for Understanding in Math. *Contemporary Educational Psychology*, 27(3), 373-391. <http://dx.doi.org/10.1006/ceps.2001.1101>
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-Set Interventions Are a Scalable Treatment for Academic Underachievement. *Psychological Science*. doi:10.1177/0956797615571017
- Paunesku, D., Goldman, D., & Dweck, C. S. (2011). *Secondary School Mindset Study*. Glasgow: The Centre for Confidence & Well-being.
- Paunesku, D., Goldman, D., & Dweck, C. S. (2011). *East Renfrewshire Growth Mindset Study*. Glasgow: The Centre for Confidence & Well-being.
- Rienzo, C., Rolfe, H., & Wilkinson, D. (2015). *Changing mindsets: Evaluation report and executive summary*. London: Education Endowment Foundation.
- Simms, V. (2016). Mathematical mindsets: unleashing students' potential through creative math, inspiring messages and innovative teaching. *Research in Mathematics Education*, 18(3), 317-320. doi:10.1080/14794802.2016.1237374
- Wheater, R., Ager, R., Burge, B., & Sizmur, J. (2014). *Achievement of 15-Year-Olds in England: PISA 2012 National Report (OECD Programme for International Student Assessment) [Revised Version, April 2014]*. London: Department for Education.

11 Transition from Primary to Secondary

What is the evidence regarding how teaching can support learners in mathematics across the transition between Key Stage 2 and Key Stage 3?

The evidence indicates a large dip in mathematical attainment as children move from primary to secondary school in England, which is accompanied by a dip in learner attitudes. There is very little evidence concerning the effectiveness of particular interventions that specifically address these dips. However, research does indicate that initiatives focused on developing shared understandings of curriculum, teaching and learning are important. Both primary and secondary teachers are likely to be more effective if they are familiar with the mathematics curriculum and teaching methods outside of their age phase. Secondary teachers need to revisit key aspects of the primary mathematics curriculum, but in ways that are engaging and relevant and not simply repetitive. Teachers' beliefs about their ability to teach appear to be particularly crucial for lower-attaining students in Key Stage 3 mathematics.

Strength of evidence: LOW

Findings

Evidence indicates a significant dip in mathematical attainment at transition. For example, in a large national study of primary attainment in England, Brown et al. (2008) found that, at the end of Year 7, a full year after the transition to secondary school, learners' performance on a test of primary numeracy was below their performance at the end of Year 6, and the impact was roughly equivalent to an ES of $d = -0.1$. (See also Galton et al., 2003.) Learners' attitudes to mathematics also decrease across the transition and continue to fall throughout Key Stage 3 (Galton et al., 2003; Zanobini & Usai, 2002).

There are a number of potential causes for the dip in attainment. Drawing on an extensive evidence base, Galton et al. (2003) found that, alongside the emotional and social adjustment to a very different school environment, there are considerable discontinuities in the curriculum, how it is taught and how learners are grouped (see also Symonds & Galton, 2014; Jansen et al., 2012). In addition, it is thought that, in Year 6, the focus on revision for national tests may result in children experiencing a narrow curriculum and a restricted range of teaching approaches, which in turn has negative implications for their mathematics learning in lower secondary (Galton et al., 2003).

Symonds & Galton's (2014) review found that secondary teachers often 'start from scratch' without reference to test results or information from primary schools (see also Galton et al., 2003). This may be compounded by untested – and, based on Galton et al.'s ORACLE studies, largely incorrect – assumptions about primary practice, which “either underestimate the demands primary teachers make on pupils ... or make assumptions about the exposure of pupils to more sophisticated forms of learning” (Galton et al., 2003, p. 26). Indeed, some studies have found that, in Year 7, tasks are at a *lower* level of challenge than learners' prior attainment in Year 6. As a result, learners may become bored or frustrated. Moreover, teachers do not always teach learners about new or more sophisticated forms of learning. Galton et al. highlight a need to place more emphasis on transition initiatives relating to

curriculum, teaching and learning, although relatively few initiatives used by schools actually do this.

Galton et al. (2003) found a range of innovative approaches to transition. They consider Integrated Learning Systems (ILS) to have potential to support learners with weaknesses in mathematics (see Technology module). However, there appears to be no evidence concerning the effectiveness of this, or any other interventions or strategies that support learners at the primary-secondary transition (McGee et al., 2003; see also the parallel Dowker review of interventions).

Secondary teachers face a dilemma. They need to revisit aspects of the primary mathematics curriculum, whilst setting an appropriate level of challenge and avoiding learner boredom or frustration. McGee et al. (2003) found that learners are likely to benefit from more task-focused instructional practices and additionally cite Midgley & Maehr's (1998) recommendation to focus on mastery, understanding and challenge. In order to do this, it is important that primary and secondary teachers are familiar with the curriculum and teaching approaches commonly used in their respective phases. Indeed, Ma's (1999) comparison of Chinese and US teachers suggests that teachers may be more effective if they are familiar with the mathematics curriculum that students have encountered in previous years and that they will encounter in later years.

Teachers' beliefs appear to be particularly crucial for lower-attaining students. In a longitudinal study of 1,329 students before and after the transition to junior high school (Key Stage 4), Midgley et al. (1989) examined the relationship between teachers' personal efficacy, their belief that their own teaching could make a difference to all learners, and learners' attitudes to mathematics. They found teachers' personal efficacy to be a strong positive influence on low attainers' attitudes to mathematics, whereas it appeared to make no difference to high attainers' attitudes.

Evidence base

We found no meta-analyses addressing the issue of transition. We draw on two systematic reviews and several single studies. As highlighted above, there is no research evidence concerning the effectiveness of specific interventions in this area.

Directness

Much of the primary research cited in the two systematic reviews has been conducted in England. Although some of the studies were conducted some time ago, there is no reason to suppose that the current situation is any different.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	
How the intervention was defined and operationalised	1	We identified no research on the effectiveness of interventions to support learners across the transition from primary to secondary.
Any reasons for possible ES inflation	3	

Any focus on particular topic areas	2	Research on transitions tends to focus on the effects on learners' attitudes and attainment in general, rather than specifically in mathematics.
Age of participants	3	

References

Systematic reviews included

Symonds, J. E., & Galton, M. (2014). Moving to the next school at age 10–14 years: an international review of psychological development at school transition. *Review of Education*, 2(1), 1-27. doi: 10.1002/rev3.3021

McGee, C., Ward, R., Gibbons, J., & Harlow, A. (2003). Transition to secondary school: A literature review. A Report to the Ministry of Education. Hamilton, University of Waikato, New Zealand.

Other references

Brown, M., Askew, M., Hodgen, J., Rhodes, V., Millett, A., Denvir, H., & William, D. (2008). Individual and cohort progression in learning numeracy ages 5-11: Results from the Leverhulme 5-year longitudinal study. In A. Dowker (Ed.), *Mathematical Difficulties: Psychology and Intervention* (pp. 85-108). Oxford: Elsevier.

Galton, M., Gray, J., Ruddock, J., Berry, M., Demetriou, H., Edwards, J., . . . Charles, M. (2003). Transfer and transitions in the middle years of schooling (7-14): Continuities and discontinuities in learning. London: DfES.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self-and task-related beliefs in mathematics during the transition to junior high school. *Journal of educational Psychology*, 81(2), 247.

Midgley, C. & Maehr, M.L. (1998). The Michigan middle school study: Report to participating schools and districts. Ann Arbor, MI: University of Michigan.

Zanobini, M., & Usai, M. C. (2002). Domain-specific self-concept and achievement motivation in the transition from primary to low middle school. *Educational Psychology*, 22(2), 203-217.

12 Teacher Knowledge and Professional Development

What is the evidence regarding the impact of teachers and their effective professional development in mathematics?

The evidence shows that the quality of teaching makes a difference to student outcomes. The quality of teaching, or instructional guidance, is important to the efficacy of almost every strategy that we have examined. The evidence also indicates that, in mathematics, teacher knowledge is a key factor in the quality of teaching. Teacher knowledge, more particularly pedagogic content knowledge (PCK), is crucial in realising the potential of mathematics curriculum resources and interventions to raise attainment. Professional development (PD) is key to raising the quality of teaching and teacher knowledge. However, evidence concerning the specific effects of PD is limited. This evidence suggests that extended PD is more likely to be effective than short courses.

Strength of evidence (Teacher knowledge): LOW

Strength of evidence (Teacher PD): LOW

Findings

Across the strategies, approaches and interventions we have examined in this review, the role of the teacher consistently comes across as a crucial, and often mediating, factor in the success of any approach. The evidence shows that the quality of teaching makes a difference to student outcomes and that a crucial factor is teacher knowledge (Coe et al., 2014).

The impact of teacher knowledge

A central component of teacher knowledge is *content knowledge* (CK). The association between teacher CK and student attainment in mathematics is well-established (see, for example, the many studies cited in Hill et al., 2005). As Coe et al. (2014, p. 18) observe, it is “intuitively obvious” that teachers need to understand the things that they teach. However, knowledge of mathematics alone appears not to be sufficient, and a great deal of research has investigated the role of what Shulman (1987) termed *pedagogical content knowledge* (PCK), as distinct from CK.

For Shulman (1987, p. 8, original emphasis), PCK concerns “subject matter knowledge *for teaching*”; notably, how a teacher translates their CK into something accessible to learners. PCK comprises, among many other factors: knowing the appropriate curriculum to teach, having a sense of learning trajectories through the subject and constituent topics, knowing the questions to ask, when and where different representations are appropriate, and the multitude of ways they may guide a learner through a problem, being confident in responding to learners’ explanations and, in so doing, recognising and addressing misconceptions.

Rowland et al. (2009) have developed Shulman’s categories into a ‘Knowledge Quartet’ to support and focus primary teachers in reflecting on what they know and do in teaching mathematics. Akin to Shulman’s CK, Rowland et al.’s quartet includes a need to make sense of foundation knowledge. Further, the quartet includes the categories of transformation, connections and contingency, all of which resonate with aspects of Shulman’s PCK, and which emphasise the importance of teachers’ knowledge going beyond the subject matter.

In an analysis of an extension study to PISA 2003 in Germany, Baumert et al. (2010) examined the different effects of mathematics teachers' CK and PCK. They defined PCK as having three dimensions: knowledge of mathematical tasks as instructional tools, knowledge of students' thinking and assessment, and knowledge of multiple representations and explanations of mathematics. Baumert et al. (2010) found that, although PCK and CK were strongly correlated, PCK was a stronger predictor of student progress than CK, after controlling for other factors. Overall, the effect size estimate of PCK for these teachers was 0.33. In addition, Baumert et al. (2010) showed that the effect of PCK was fully mediated by three factors – the choice and enactment of tasks, the alignment of instruction to the curriculum, and the adaptation of instruction for learners¹⁷ – whereas CK was only mediated by alignment of instruction to the curriculum. In other words, the quality of learning opportunities is largely determined by PCK.

In a study of first and third grade (Y2 and Y4) teachers, Hill et al. (2005) investigated the effects of mathematical knowledge for teaching (MKT) on attainment. Defining MKT as an amalgam of knowledge, including PCK, Hill et al. (2005) found that MKT had an effect equivalent to more than a month's learning (based on a comparison of teachers with high and low knowledge); an effect of roughly similar size to the effects of student SES or ethnicity.

In a study of effective teaching of numeracy focused on primary teaching in England, Askew et al. (1997) focused on effectiveness as defined by learning gains over the course of a year. Based on a sample of 72 teachers, they found that highly effective teachers used teaching approaches that emphasised connections between different areas of mathematics and believed that learners learn mathematics by being challenged to think, through explaining, listening and problem solving. Being highly effective was not associated with having an A-level or degree in mathematics. Highly effective teachers were more likely than other teachers to have participated in mathematics-specific PD over an extended period.

These findings do not mean that a teacher *only* needs PCK to be effective, or that CK is unimportant. As Baumert et al. (2010) argue, PCK is inconceivable without CK. Moreover, their findings indicate that it is not possible to compensate for weak CK by focusing on PCK in teacher education. In short, CK is a necessary but not sufficient condition for high-quality teaching.

The impact of professional development

Professional development (PD) is key to raising the quality of teaching and teacher knowledge. Elsewhere in this review, we have highlighted the importance of PD to the effectiveness of a number of strategies. Indeed, it is difficult to imagine how teachers could learn how to implement many of the strategies referred to in this review without some kind of PD. However, although many interventions in this review involve PD of some kind, we found little evidence about the effectiveness of PD itself. Much of the evidence of effectiveness draws on teacher self-report (e.g., Back et al., 2009), despite the fact that teacher perceptions of changes to their practice

¹⁷Baumert et al (2010) refer to the three factors as (i) *cognitive activation* of tasks: how the teacher supports learners in developing problem solving strategies, understanding methods and constructing connections between and within mathematical topics; (ii) *alignment of instruction* to the Grade 10 curriculum; and (iii) *individual learning support*, the extent to which the teacher adapted explanations, responded constructively to errors, set an appropriate pace and whether interactions were respectful and caring.

have been shown in many studies to be unreliable (e.g., Lortie & Spillane, 1999). Timperley et al. (2007) found that training sessions of a day or less, the dominant model of PD, can be useful for more straightforward aims, for example the transmission of new educational policies or strategies such as curriculum specification changes, but are unlikely to enable teachers to transform the quality of their instructional practice or their pedagogical content knowledge. Yoon et al.'s (2007) review of PD found that PD of 14 hours or more was associated with modest, statistically significant student gains in attainment, whereas anything of shorter duration produced no gains. Yoon et al. found that PD of substantial duration (an average of 49 hours) was associated with an average student gain of $d=0.54$. However, this was based on only nine studies, of which four were in mathematics, and we note that Yoon et al. comment that rigorous studies are needed to better understand the effect of duration alongside other characteristics, such as intensity.

There are few robust experimental studies that isolate the effects of PD in mathematics. Gersten et al.'s (2014) review found only five studies focusing on different approaches to PD. Of these five, only two showed significant positive effects on learners' attainment, whilst two approaches showed no effect on learner attainment and the fifth had limited effects. Gersten et al.'s study adds weight to Yoon et al.'s call for more research. It is widely considered that significant professional change takes a considerable amount of time to develop, with many suggesting a period of up to two years (e.g., Adey et al., 2004; Clarke, 2004). If these judgments are correct, the impact of PD on learner attainment is likely to take some time to develop. Hence, there is a need for longitudinal studies of the impact of PD.

Evidence base

We have drawn on one meta-analysis considered to be of medium methodological quality. This was supplemented by three research syntheses and a range of other literature including seminal works in the area. The consistency in the commentary arising from these studies is strong, although the evidence base is weak.

Overall, the number of robust experimental studies into effective PD programmes in mathematics is very small. It should be noted that the limited number of studies may be accounted for by the application of the strict WWC evidence standards (version 2.1) to study inclusion; the authors passed 32 studies through 3 previous screening phases before the WWC standards reduced the included studies to five.

Given the recognised importance of the quality of teaching for learner outcomes, and the limited number of studies of PD in mathematics, this is an important area for further research.

Directness

One study included within our literature (Baumert et al., 2010) was a study of 15-year-olds, which we acknowledge as sitting outside of our 9-14 remit. However, this study is nonetheless applicable to our target age group, as the focus is on secondary mathematics teachers generally, rather than age-specific topics or approaches, and covered the full attainment range in German Grade 10 classes.

Threat to directness	Grade	Notes
Where and when the studies were carried out	3	Studies are upto date and generally reflect current educational policy. Studies were set in a range of contexts, including England.
How the intervention was defined and operationalised	3	
Any reasons for possible ES inflation	3	
Any focus on particular topic areas	2	
Age of participants	2	

Overview of Effects

Study	Effect size	Notes
<i>Effects of PD on student attainment</i>		
Yoon et al. (2007); <i>overall effect of PD on attainment, all included studies</i>	0.54	Substantial PD (an average of 49 hours) was associated with an average student gain of $d=0.54$. The 20 ESs across the nine studies ranged from -0.53 to 2.39 .
Yoon et al. (2007); <i>overall effect of PD on mathematics attainment</i>	0.57	The 6 ESs across the four studies ranged from -0.53 to 2.39 . It should be noted that both of these extremes came from the same study (Saxe et al., 2001). The other ESs were 0.26 , 0.41 , 0.41 and 0.5 . The average ES of 0.57 for mathematics attainment compares with average ESs of 0.51 and 0.53 for science and for reading/English/language arts respectively.

Meta-analyses

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>

Meta-analyses Excluded

Salinas, A. (2010). Investing in our teachers: What focus of professional development leads to the highest student gains in mathematics

achievement? PhD Thesis, University of Miami [ES for equity and PCK too large]

Systematic reviews

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). *What makes great teaching? Review of the underpinning research*. London: The Sutton Trust.

Gersten, R., Taylor, M. J., Keys, T. D., Rolfhus, E., & Newman-Gonchar, R. (2014). *Summary of research on the effectiveness of math professional development approaches. (REL 2014-010)*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development. Best evidence synthesis iteration (BES)*. Wellington, New Zealand: Ministry of Education.

Other references

Adey, P., Hewitt, G., Hewitt, J., & Landau, N. (2004). *The professional development of teachers: Practice and theory*. Dordrecht: Kluwer.

Askew, M., Brown, M., Rhodes, V., Johnson, D. C., & Wiliam, D. (1997). *Effective teachers of numeracy*. London: King's College.

Back, J., De Geest, E., Hirst, C., Joubert, M., & Sutherland, R. (2009). *Final Report: Researching Effective CPD in Mathematics Education (RECME)*. Sheffield, UK: National Centre for Excellence in Teaching Mathematics.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133-180. doi:10.3102/0002831209345157

Clarke, D. M. (1994). Ten key principles from research for the professional development of mathematics teachers. In D. B. Aichele & A. F. Coxford (Eds.), *Professional development for teachers of mathematics: The 1994 Yearbook of the National Council of Teachers of Mathematics* (pp. 37-48). Reston, VA: National Council of Teachers of Mathematics.

Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Rowland, T., Turner, F., Thwaites, A. & Huckstep, P. (2009) *Developing Primary Mathematics Teaching: Reflecting on Practice with the Knowledge Quartet*. London: Sage.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *Journal of*

Curriculum Studies, 31(2), 143 - 175.

13 References

* primary meta-analyses; † secondary meta-analyses

- Adey, P. S., & Shayer, M. (1994). Really raising standards. London: Routledge.
- Adey, P., Hewitt, G., Hewitt, J., & Landau, N. (2004). The professional development of teachers: Practice and theory. Dordrecht: Kluwer.
- Adhami, M. (2002). Cognitive acceleration in mathematics education in years 5 and 6: Problems and challenges. In M. Shayer & P. S. Adey (Eds.), *Learning Intelligence: Cognitive Acceleration across the Curriculum from 5 to 15 years* (pp. 98-117). Buckingham: Open University Press.
- Advisory Committee on Mathematics Education [ACME]. (2011). *Mathematical Needs: Mathematics in the workplace and in Higher Education*. London: Royal Society.
- Advisory Committee on Mathematics Education [ACME]. (2016). *Problem solving in mathematics: realising the vision through better assessment*. London: Royal Society.
- Ahmed, A. (1987). *Better mathematics: A curriculum development study based on the low attainers in mathematics project*. London: HM Stationery Office.
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198. doi: 10.1016/j.learninstruc.2006.03.001
- Alcock, L., Ansari, D., Batchelor, S., Bisson, M.-J., De Smedt, B., Gilmore, C., . . . Weber, K. (2016). Challenges in mathematical cognition: a collaboratively derived research agenda. *Challenges in Mathematical Cognition*, 1(2), 20-41.
- Alexander, R. J. (2017). *Towards dialogic teaching: Rethinking classroom talk*. (5th edition) York: Dialogos.
- Alexander, R.J. (ed) (2010) *Children, their World, their Education: final report and recommendations of the Cambridge Primary Review*. London: Routledge.
- Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/Pàngarau: Best evidence synthesis iteration*. Wellington, New Zealand: Ministry of Education.
- Askew, M., Brown, M., Rhodes, V., Johnson, D. C., & Wiliam, D. (1997). *Effective teachers of numeracy*. London: King's College.
- Askew, M., Hodgen, J., Hossain, S., & Bretscher, N. (2010). *Values and variables: Mathematics education in high-performing countries*. London: Nuffield Foundation.
- * Athappilly, K., Smidchens, U., & Kofel, J. W. (1983). A computer-based meta-analysis of the effects of modern mathematics in comparison with traditional mathematics. *Educational Evaluation and Policy Analysis*, 485-493.
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research. *Review of Educational Research*, 70(2), 181-214. doi:10.3102/00346543070002181
- Back, J., De Geest, E., Hirst, C., Joubert, M., & Sutherland, R. (2009). *Final Report: Researching Effective CPD in Mathematics Education (RECME)*. Sheffield, UK: National Centre for Excellence in Teaching Mathematics.

- Baines, E., Blatchford, P., Kutnick, P., Chowne, A., Ota, C., & Berdondini, L. (2014). *Promoting Effective Group Work in the Primary Classroom: A handbook for teachers and practitioners*. London: Routledge.
- * Baker, S., Gersten, R., & Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, 103(1), 51-73. doi:10.1086/499715
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of personality and social psychology*, 41(3), 586.
- † Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61(2), 213-238. doi:10.3102/00346543061002213
- Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. J. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 843-908). Greenwich, CT: Information Age Publishing.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133-180. doi:10.3102/0002831209345157
- * Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education: Innovations and Research*, 12(5/6), 23.
- Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296-233). New York: Macmillan.
- Bierhoff, H. (1996). *Laying the foundations of numeracy: A comparison of primary school textbooks in Britain, Germany and Switzerland*. London: National Institute for Economic and Social Research.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bloom, B. S. (1968, May). Mastery learning. In *Evaluation comment* (Vol. 1, No. 2). Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs.
- Boaler, J. (2013). Ability and Mathematics: the mindset revolution that is reshaping education. *Forum*, 55(1), 143-152.
- * Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39-55.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience, and School* (Expanded Edition). Washington, DC: National Academies Press.
- Brooks, G., Pahl, K., Pollard, A., & Rees, F. (2008). *Effective and inclusive practices in family literacy, language and numeracy: a review of programmes and practice in the UK and internationally*. Reading: CfBT.
- * Browder, D. M., Spooner, F., Ahlgrim-DeLzell, L., Harris, A. A., & Wakemanxya, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional Children*, 74(4), 407-432.

- Brown, A. L., & Campione, J. C. (1990). Interactive learning environments and the teaching of science and mathematics. In M. Gardner et al. (Eds.), *Toward a scientific practice of science education*. Hillsdale, NJ: Erlbaum.
- Brown, M. (1989). Graded Assessment and Learning Hierarchies in Mathematics: An alternative view. *British Educational Research Journal*, 15(2), 121 - 128. doi:10.1080/0141192890150202
- Brown, M. (1996). The context of the research: the evolution of the national curriculum for mathematics. In D. C. Johnson & A. Millett (Eds.), *Implementing the Mathematics National Curriculum: policy, politics and practice* (pp. 1-28). London: Paul Chapman Publishing.
- Brown, M. (Ed.) (1992). *Graded Assessment in Mathematics (GAIM) topic criteria handbook*. Walton on Thames: Nelson.
- Brown, M., Askew, M., Baker, D., Denvir, H., & Millett, A. (1998). Is the National Numeracy Strategy research-based? *British Journal of Educational Studies*, 46(4), 362-385.
- Brown, M., Askew, M., Hodgen, J., Rhodes, V., Millett, A., Denvir, H., & Wiliam, D. (2008). Individual and cohort progression in learning numeracy ages 5-11: Results from the Leverhulme 5-year longitudinal study. In A. Dowker (Ed.), *Mathematical Difficulties: Psychology and Intervention* (pp. 85-108). Oxford: Elsevier.
- Brown, M., Askew, M., Millett, A., & Rhodes, V. (2003). The key role of educational research in the development and evaluation of the National Numeracy Strategy. *British Educational Research Journal*, 29(5), 655-672.
- Brown, M., Blondel, E., Simon, S., & Black, P. (1995). Progression in measuring. *Research Papers in Education*, 10(2), 143-170. doi:10.1080/0267152950100202
- Brown, M., Brown, P., & Bibby, T. (2008). "I would rather die": Attitudes of 16 year-olds towards their future participation in mathematics. *Research in Mathematics Education*, 10(1), 3-18.
- Bryant, P. (2009). Paper 5: Understanding space and its representation in mathematics. In T. Nunes, P. Bryant, & A. Watson (Eds.), *Key understandings in mathematics learning*. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 4 December 2017.
- Bryant, P., & Nunes, T. (2012). *Children's understanding of probability: A literature review*. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 4 December 2017.
- *Burns, M. K., Coddington, R. S., Boice, C. H., & 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(1), 69.
- Canadian Council on Learning (2009). *A systematic review of literature examining the impact of homework on academic achievement*. Toronto: Canadian Council on Learning.
- *Capar, G., & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory & Practice*, 2, 553-559.
- Cara, O., & Brooks, G. (2012). *Evidence of the Wider Benefits of Family Learning: A Scoping Review*. BIS Research Paper Number 93. London: Department for Business, Innovation and Skills.

- * 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(2), 380. doi: 10.1037/a0031084
- Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann.
- * Chan, K. K., & Leung, S. W. (2014). Dynamic geometry software improves mathematical achievement: Systematic review and meta-analysis. *Journal of Educational Computing Research*, 51(3), 311-325.
- * Chauhan, S. (2017). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education*, 105, 14-30.
- *† Chen, H. (2004). The efficacy of mathematics interventions for students with learning disabilities: A meta-analysis. (Order No. 3157959, The University of Iowa). Retrieved from <http://search.proquest.com/docview/305196240?accountid=14533>. (305196240).
- Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology*, 107(3), 689.
- * 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.
- Chinn, S. (2009). Mathematics anxiety in secondary students in England. *Dyslexia*, 15(1), 61-68.
- Christiansen, B., & Walther, G. (1986). Task and activity. In B. Christiansen, A. G. Howson, & M. Otte (Eds.), *Perspectives in mathematics education* (pp. 243–307). Dordrecht: Reidel.
- Churches, R. (2016). *Closing the gap: test and learn*. London: Department for Education.
- Clarke, D. M. (1994). Ten key principles from research for the professional development of mathematics teachers. In D. B. Aichele & A. F. Coxford (Eds.), *Professional development for teachers of mathematics: The 1994 Yearbook of the National Council of Teachers of Mathematics* (pp. 37-48). Reston, VA: National Council of Teachers of Mathematics.
- Clarke, D., Sullivan, P., Cheeseman, J., & Clarke, B. (2000). The Early Numeracy Research Project: Developing a framework for describing early numeracy learning. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000 (Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia)* (pp. 180-187). Freemantle, Western Australia: MERGA.
- Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 420-464). New York: Macmillan.
- Clements, D. H., & Samara, J. (2014). *Learning and Teaching Early Math: The Learning Trajectories Approach*. Oxford: Routledge.
- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical thinking and learning*, 6(2), 81-89.
- Cockcroft, W. H. (1982). *Mathematics counts*. London: HMSO.

- * Coddling, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice*, 26(1), 36-47.
- Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). What makes great teaching? Review of the underpinning research. London: The Sutton Trust.
- * Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational Outcomes of Tutoring: A Meta-analysis of Findings. *American Educational Research Journal*, 19(2), 237-248.
- Confrey, J., Maloney, A. P., Nguyen, K. H., Mojica, G., & Myers, M. (2009). Equipartitioning / splitting as a foundation of rational number reasoning using learning trajectories. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 345-352). Thessaloniki, Greece: PME.
- Conway, P. (2005). International trends in post-primary mathematics education. Research report commissioned by the National Council for Curriculum and Assessment (NCCA) Retrieved from <http://www.ncca.ie/uploadedfiles/mathsreview/intpaperoct.pdf>
- Cooper, B., & Dunne, M. (2000). Assessing children's mathematical knowledge: Social class, sex and problem-solving. Buckingham: Open University Press.
- * Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of educational research*, 76(1), 1-62.
- * Crawford, S. T. (2011). *Meta-Analysis of the Impact of After-School Programs on Students Reading and Mathematics Performance*. (ProQuest UMI 3486475 EdD), University Of North Texas.
- Daro, P., Mosher, F. A., & Corcoran, T. (2011). Learning trajectories in mathematics: a foundation for standards, curriculum, assessment, and instruction. Philadelphia, PA: Consortium for Policy Research in Education (CPRE).
- De Lange, J. (1996). Using and applying mathematics in education. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 49-98). Dordrecht, The Netherlands: Kluwer.
- de Lange, J. (1999). Framework for classroom assessment in mathematics. Utrecht: Freudenthal Institute & National Center for Improving Student Learning and Achievement in Mathematics and Science.
- * Demir, S., & Basol, G. (2014). Effectiveness of Computer-Assisted Mathematics Education (CAME) over Academic Achievement: A Meta-Analysis Study. *Educational Sciences: Theory and Practice*, 14(5), 2026-2035.
- * Dennis, M. S., Sharp, E., Chovanes, J., Thomas, A., Burns, R. M., Custer, B., & Park, J. (2016). A Meta-Analysis of Empirical Research on Teaching Students with Mathematics Learning Difficulties. *Learning Disabilities Research & Practice*, 31(3), 156-168.
- Denvir, B., & Brown, M. (1986). Understanding number concepts in low attaining 7-9 year olds. Part II: The teaching studies. *Educational Studies in Mathematics*, 17, 143-164.

- DfEE (1999). The National Numeracy Strategy: Framework for teaching mathematics from Reception to Year 6. London: Department for Education and Employment.
- DfEE (2001). Framework for teaching mathematics Years 7, 8 and 9. London: Department for Education and Employment.
- DfEE. (1998). Framework for teaching mathematics: Reception to Year 6 (September ed.). London: Department for Education and Employment.
- Dickson, L., Brown, M., & Gibson, O. (1984). Children learning mathematics: A teacher's guide to recent research. Eastbourne: Holt, Rinehart & Winston Ltd.
- * Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231-264.
- * Domino, J. (2010). The Effects of Physical Manipulatives on Achievement in Mathematics in Grades K-6: A Meta-Analysis. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.
- * Donker, A. S., De Boer, H., Kostons, D., van Ewijk, C. D., & Van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1-26. DOI: 10.1016/j.edurev.2013.11.002
- Donovan, M. S., & Bransford, J. D. (2005). How students learn: Mathematics in the classroom. Washington, DC: The National Academies Press.
- Dowker, A. (2014). Young children's use of derived fact strategies in addition and subtraction. *Frontiers in Human Neuroscience*, 7. doi:10.3389/fnhum.2013.00924
- Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics Anxiety: What Have We Learned in 60 Years? *Frontiers in Psychology*, 7(508). doi:10.3389/fpsyg.2016.00508.
- Dugan, J.J. (2007). A systematic review of interventions in secondary mathematics with at-risk students: Mapping the Literature Doctoral dissertation Colorado State University, Fort Collins Colorado. ProQuest UMI 3266389
- * Durkin, K. (2011). The self-explanation effect when learning mathematics: A meta-analysis. Presented at the Society for Research on Educational Effectiveness 2011, Available online [accessed 13th March 2017]: <http://eric.ed.gov/?id=ED518041>.
- Education Endowment Foundation (2017) Teaching & Learning Toolkit: Parental engagement. London: EEF.
- † Education Endowment Foundation (2017) Teaching & Learning Toolkit: Setting or streaming. London: EEF.
- * Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, 34, 433-463.
- Ellis, A. K., Bond, J. B., & Denton, D. W. (2012). An analytical literature review of the effects of metacognitive teaching strategies in primary and secondary student populations. *Asia Pacific Journal of Educational Development (APJED)*, 1(1), 9-23.
- Ellis, A. K., Denton, D. W., & Bond, J. B. (2014). An analysis of research on metacognitive teaching strategies. *Procedia-Social and Behavioral Sciences*, 116, 4015-4024.

- Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: development status and directions. *ZDM Mathematics Education*, 45(5), 633-646.
- Foster, C., & Inglis, M. (2017). Teachers' appraisals of adjectives relating to mathematics tasks. *Educational Studies in Mathematics*, 95(3), 283–301. <https://doi.org/10.1007/s10649-017-9750-y>
- †Friso-van den Bos, I., van der Ven, S. H., Kroesbergen, E. H., & van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational research review*, 10, 29-44.
- Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). Teaching math to young children: A practice guide (NCEE 2014-4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Fuson, K. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 243-295). New York: Macmillan.
- Galton, M., Gray, J., Ruddock, J., Berry, M., Demetriou, H., Edwards, J., . . . Charles, M. (2003). Transfer and transitions in the middle years of schooling (7-14): Continuities and discontinuities in learning. London: DfES.
- Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, 5(1), 3-57. doi:10.1002/rev3.3077
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools (NCEE 2009-4060). Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Educational Sciences, US Department of Education.
- *Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202-1242
- Gersten, R., Taylor, M. J., Keys, T. D., Rolfhus, E., & Newman-Gonchar, R. (2014). Summary of research on the effectiveness of math professional development approaches. (REL 2014–010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.
- Gersten, R., Woodward, J., & Darch, C. (1986). Direct instruction: A research-based approach to curriculum design and teaching. *Exceptional Children*, 53(1), 17-31.
- Gilmore, C.K., Keeble, S., Richardson, S., Cragg, L. (Forthcoming) The Interaction of Procedural Skill, Conceptual Understanding and Working Memory in Early Mathematics Achievement, *Journal of Numerical Cognition*.
- Gravemeijer, K. P. E. (1994). *Developing Realistic Mathematics Education*. Utrecht: CD-β Press / Freudenthal Institute.
- Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for research in Mathematics Education*, 116-140.

- Greany, T., Barnes, I., Mostafa, T., Pesniero, N., & Swenson, C. (2016). Trends in Maths and Science Study (TIMSS): National Report for England. London: Department for Education.
- Greer, B. (1994). Extending the Meaning of Multiplication and Division. In G. Harel & J. Confrey (Eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics* (pp. 61-85): SUNY Press.
- Griffiths, R., Back, J., & Gifford, S. (2016). *Making numbers: Using manipulatives to teach arithmetic*. Oxford: Oxford University Press.
- * Guskey, T. R., & Pigott, T. D. (1988). Research on Group-Based Mastery Learning Programs: A Meta-Analysis. *The Journal of Educational Research*, 81(4), 197-216. doi:10.1080/00220671.1988.10885824
- * Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin*, 89(642), 24-46.
- Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, 28(4), 567-590.
- Hart, K. (Ed.) (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- Hart, K. M., Brown, M. L., Kuchemann, D. E., Kerslake, D., Ruddock, G., & McCartney, M. (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.
- * Hartley, S. S. (1977) *Meta-Analysis of the Effects of Individually Paced Instruction In Mathematics*. Doctoral dissertation University of Colorado at Boulder.
- † Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- * Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(1), 33-46.
- * Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. *Journal for Research in Mathematics Education*, 242-273.
- * Hembree, R., & Dessart, D. J. (1986). Effects of hand-held calculators in precollege mathematics education: A meta-analysis. *Journal for research in mathematics education*, 17(2), 83-99.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.
- Higgins, S., & Katsipatakis, M. (2016). Communicating comparative findings from meta-analysis in educational research: some examples and suggestions. *International Journal of Research & Method in Education*, 39(3), 237-254. doi:10.1080/1743727X.2016.1166486
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- † Higgins, S., Katsipatakis, M., Villanueva-Aguilera, A. B. V., Coleman, R., Henderson, P., Major, L. E., Coe, R., & Mason, D. (2016). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Feedback*. London: Education Endowment Foundation.
<https://educationendowmentfoundation.org.uk/resources/teachinglearning-toolkit>

- Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- * Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: a meta-analytic assessment of the strategies that promote achievement. *Developmental psychology*, 45(3), 740.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational psychologist*, 42(2), 99-107.
- HMI. (1985). Mathematics from 5 to 16. HMI Series: Curriculum Matters No. 3. London: HMSO.
- Hodgen, J. (2012). Computers good, calculators bad. In P. Adey & J. Dillon (Eds.), *Bad Education: Debunking educational myths*. Maidenhead: Open University Press.
- Hodgen, J., & Marks, R. (2013). *The Employment Equation: Why our young people need more maths for today's jobs*. London: The Sutton Trust.
- Hodgen, J., Brown, M., Küchemann, D., & Coe, R. (2011). Why have educational standards changed so little over time: The case of school mathematics in England. Paper presented at the British Educational Research Association (BERA) Annual Conference, Institute of Education, University of London.
- Hodgen, J., Küchemann, D., & Brown, M. (2010). Textbooks for the teaching of algebra in lower secondary school: are they informed by research? *Pedagogies*, 5(3), 187-201. doi:10.1080/1554480X.2013.739275
- * Holmes, A. B. (2013). Effects of Manipulative Use on PK-12 Mathematics Achievement: A Meta-Analysis. Society for Research on Educational Effectiveness.
- Howson, G. (2013). The development of mathematics textbooks: historical reflections from a personal perspective. *ZDM Mathematics Education*, 45(5), 647-658.
- * Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A Meta-Analysis of Algebra Interventions for Learners with Disabilities and Struggling Learners. *Journal of the International Association of Special Education*, 15(1).
- Jacob, R., & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement. *Review of Educational Research*, 85(4), 512-552. doi:10.3102/0034654314561338
- * Jacobse, A. E., & Harskamp, E. (2011). A meta-Analysis of the Effects of instructional interventions on students' mathematics achievement. Groningen: GION, Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling, Rijksuniversiteit Groningen.
- Jaworski, B. (1994). *Investigating mathematics teaching: A constructivist enquiry*. London: Falmer Press.
- Jerrim, J., & Shure, N. (2016). *Achievement of 15-Year-Olds in England: PISA 2015 National Report*. London: Department for Education.
- Jerrim, J., & Vignoles (2015). *Mathematics Mastery: Overarching Summary Report*. Education Endowment Foundation.
- Jones, G. A., Langrall, C. W., & Monney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. J. Lester (Ed.), *Second*

- handbook of Research on mathematics teaching and learning (pp. 909-955). Greenwich, CT: Information Age Publishing.
- Jones, I., & Pratt, D. (2012). A Substituting Meaning for the Equals Sign in Arithmetic Notating Tasks. *Journal for Research in Mathematics Education*, 43(1), 2-33.
- Kaput, J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5-17). New York, NY: Lawrence Erlbaum.
- Keller, F. S. (1968). "Good-bye, teacher..." *Journal of Applied Behavioral Analysis*, 1, 79-89.
- Kieran, C. (2004). The core of algebra: Reflections on its main activities. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The future of the teaching and learning of algebra: The 12th ICMI Study* (pp. 21-33). Dordrecht, The Netherlands: Kluwer.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding It Up: Helping Children Learn Mathematics* (Prepared by the Mathematics Learning Study Committee, National Research Council). Washington DC: The National Academies Press.
- * Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28-37. doi:10.1111/j.1745-3992.2011.00220.x
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.
- † Kluger, A. N., & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), 254-284.
- Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal*, 40(1), 281-310.
- * Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs a meta-analysis. *Remedial and special education*, 24(2), 97-114.
- Küchemann, D. E. (1981). *The understanding of generalised arithmetic (algebra) by secondary school children* (PhD thesis). Chelsea College, University of London.
- * Kuchler, J. M. (1998) *The effectiveness of using computers to teach secondary school (grades 6-12) mathematics: A meta-analysis*. Ph.D. thesis, University of Massachusetts Lowell.
- * Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of Mastery Learning Programs: A Meta-Analysis. *Review of Educational Research*, 60(2), 265-299. doi:doi:10.3102/00346543060002265
- Kyriacou, C., & Issitt, J. (2008). *What Characterises Effective Teacher-initiated Teacher-pupil Dialogue to Promote Conceptual Understanding in Mathematics Lessons in England in Key Stages 2 and 3: A Systematic Review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Lai, E. R. (2011). *Metacognition: A literature review*. Always learning: Pearson research report.

- Lamon, S. J. (2007). Rational Numbers and Proportional Reasoning: Toward a Theoretical Framework for Research. In F. K. J. Lester (Ed.), *Second handbook of Research on mathematics teaching and learning* (pp. 629-667). Greenwich, CT: Information Age Publishing.
- * Lazowski, R. A., & Hulleman, C. S. (2016). Motivation Interventions in Education. *Review of Educational Research*, 86(2), 602-640. doi:doi:10.3102/0034654315617832
- * Lee, D. S. (2000). A meta-analysis of mathematics interventions reported for 1971-1998 on the mathematics achievement of students identified with learning disabilities and students identified as low achieving. Doctoral Thesis, University of Oregon ProQuest UMI 9963449
- * LeNoir, P. (1989). The effects of manipulatives in mathematics instruction in grades K-college: A meta-analysis of thirty years of research. Doctoral Thesis, North Carolina State University at Raleigh, ProQuest UMI 8918109.
- * Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, 107(2), 558-579.
- Levin, S. W. (1998). Fractions and division: Research conceptualizations, textbook presentations, and student performances (Doctoral dissertation, University of Chicago, 1998). *Dissertation Abstracts International*, 59, 1089A.
- * Li, Q. and Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, 22(3), 215-243.
- Lloyd, C., Edovald, T., Morris, S., Kiss, Z., Skipp, A., & Haywood, S. (2015) *Durham Shared Maths Project: Evaluation Report and Executive Summary*. London: EEF
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, 66, 423-458.
- Lubienski, S. (2000). Problem solving as a means toward mathematics for all: An exploratory look through a class lens. *Journal for Research in Mathematics Education*, 31(4), 454-482.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- † Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, 30(5), 520-540.
- * Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: a meta-analysis. *Journal for Research In Mathematics Education*, 28(1), 26-47.
- Machin, S., & McNally, S. (2009). *The Three Rs: What Scope is There for Literacy and Numeracy Policies to Raise Pupil Achievement?*
- Marks, R. (2016). *Ability grouping in primary schools: case studies and critical debates*. Northwich: Critical Publishing

- Mason, J. (2008). Being mathematical with and in front of learners: Attention, awareness, and attitude as sources of difference between teacher educators, teachers and learners. In B. Jaworski (Vol. Ed.) & T. Wood (Series Ed.), *Handbook of mathematics teacher education* (Vol. 4): The mathematics teacher educator as a developing professional (pp. 31–56). Rotterdam, Netherlands: Sense.
- Mason, J., & Johnston-Wilder, S. (2006). *Designing and using mathematical tasks*. Tarquin Publications.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14.
- McGee, C., Ward, R., Gibbons, J., & Harlow, A. (2003). *Transition to secondary school: A literature review. A Report to the Ministry of Education*. Hamilton, University of Waikato, New Zealand.
- McGowen, M. A., & Tall, D. O. (2010). Metaphor or Met-Before? The effects of previous experience on practice and theory of learning mathematics. *The Journal of Mathematical Behavior*, 29(3), 169-179.
- McMullen, F., & Madelaine, A. (2014). Why is there so much resistance to Direct Instruction? *Australian Journal of Learning Difficulties*, 19(2), 137-151.
- * Melby-Lervåg, M. m., & Hulme, C. c. (2013). Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology*, 49(2), 270-291
- * Methe, S. A., Kilgus, S. P., Neiman, C., & Riley-Tillman, T. C. (2012). Meta-analysis of interventions for basic mathematics computation in single-case research. *Journal of Behavioral Education*, 21(3), 230-253.
- Middleton, J. A., & Spanias, P. A. (1999). Motivation for achievement in mathematics: Findings, generalizations and criticisms of the research. *Journal for Research in Mathematics Education*, 30(1), 65-88.
- Middleton, M. J., & Midgley, C. (2002). Beyond Motivation: Middle School Students' Perceptions of Press for Understanding in Math. *Contemporary Educational Psychology*, 27(3), 373-391.
<http://dx.doi.org/10.1006/ceps.2001.1101>
- Midgley, C. & Maehr, M.L. (1998). *The Michigan middle school study: Report to participating schools and districts*. Ann Arbor, MI: University of Michigan.
- Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self-and task-related beliefs in mathematics during the transition to junior high school. *Journal of educational Psychology*, 81(2), 247.
- Muenks, K., & Miele, D. B. (Forthcoming, online first). Students' Thinking About Effort and Ability. *Review of Educational Research*, 0034654316689328. doi: 10.3102/0034654316689328
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA / Amsterdam: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College / International Association for the Evaluation of Educational Achievement (IEA).
- Mun Ling, L., & Marton, F. (2011). Towards a science of the art of teaching: Using variation theory as a guiding principle of pedagogical design. *International Journal for Lesson and Learning Studies*, 1(1), 7-22.
- Nardi, E., & Steward, S. (2003). Is mathematics T.I.R.E.D? A profile of quiet disaffection in the secondary mathematics classroom. *British Educational Research Journal*, 29(3), 345-367.

- National Council of Teachers of Mathematics (NCTM) (2000). Principles and Standards for School Mathematics. Reston, VA: NCTM.
- NCETM (2016). The Essence of Maths Teaching for Mastery. Retrieved from <https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf>
- Newton, D. P., & Newton, L. D. (2007). Could elementary mathematics textbooks help give attention to reasons in the classroom? *Educational Studies in Mathematics*, 64(1), 69-84.
- Ng, S. F., & Lee, K. (2009). The Model Method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, 40(3), 282-313.
- Nunes, T., & Bryant, P. (2009). Paper 3: Understanding rational numbers and intensive quantities. In T. Nunes, P. Bryant, & A. Watson (Eds.), *Key understandings in mathematics learning*. London: Nuffield Foundation. Available from www.nuffieldfoundation.org, accessed 9 October 2009.
- Nunes, T., Bryant, P., & Watson, A. (2009). *Key understandings in mathematics learning*. London: Nuffield Foundation.
- Nunes, T., Bryant, P., Barros, R., & Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. *British Journal of Educational Psychology*, 82(1), 136-156. doi:10.1111/j.2044-8279.2011.02033.x
- * Nunnery, J.A., Chappell, S. & Arnold, P. (2013). A meta-analysis of a cooperative learning model's effects on student achievement in mathematics. *Cypriot Journal of Educational Sciences*, 8(1), 34-48.
- OECD. (2004). Program for International Student Assessment. Learning for tomorrow's world: First results from PISA 2003 (OECD Publications No. 53799 2004). Paris: Organisation for Economic Cooperation and Development.
- * Othman, N. (1996). The effects of cooperative learning and traditional mathematics instruction in grade K-12: A meta-analysis of findings. Doctoral Thesis, West Virginia University. ProQuest UMI 9716375
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1), 1-4.
- * Paschal, R. A., Weinstein, T., & Walberg, H. J. W. (1984). The effects of homework on learning: A quantitative synthesis. *The Journal of Educational Research*, 78(2), 97-104.
- * Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent Involvement in Homework: A Research Synthesis. *Review of Educational Research*, 78(4), 1039-1101.
- Paunesku, D., Goldman, D., & Dweck, C. S. (2011). East Renfrewshire Growth Mindset Study. Glasgow: The Centre for Confidence & Well-being.
- Paunesku, D., Goldman, D., & Dweck, C. S. (2011). Secondary School Mindset Study. Glasgow: The Centre for Confidence & Well-being.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-Set Interventions Are a Scalable Treatment for Academic Underachievement. *Psychological Science*. doi:10.1177/0956797615571017
- Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory

- domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, 108(4), 455-473. doi:10.1037/edu0000079
- Pirie, S. E. B., & Kieren, T. E. (1994). Growth in mathematical understanding: How can we characterise it and how can we represent it? *Educational Studies in Mathematics*, 26, 165 - 190.
- * Preston, J. A. (2007). Student-centered versus teacher-centered mathematics instruction: A meta-analysis. Doctoral Thesis, Indiana University of Pennsylvania. ProQuest UMI 3289778.
- * Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of Instructional Improvement in Algebra A Systematic Review and Meta-Analysis. *Review of Educational Research*, 80(3), 372-400.
- Rathé, S., Torbeyns, J., Hannula-Sormunen, M., De Smedt, B., & Verschaffe, L. (2016). Spontaneous Focusing On Numerosity: A Review of Recent Research. *Mediterranean Journal for Research in Mathematics Education*.
- Reynolds, D., & Muijs, D. (1999). The effective teaching of mathematics: A review of research. *School Leadership & Management*, 19(3), 273-288.
- Rienzo, C., Rolfe, H., & Wilkinson, D. (2015). Changing mindsets: Evaluation report and executive summary. London: Education Endowment Foundation.
- * Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM*, 1-13.
- Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, 27(4), 587-597.
- * Rohrbeck, C., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240-257.
- Rosenshine, B. (2008). Five meanings of direct instruction. Center on Innovation & Improvement, Lincoln.
- Rosenshine, B. (2012). Principles of instruction: research-based strategies that all teachers should know. *American Educator*, 36(1), 12.
- * Rosli, R., Capraro, M. M., & Capraro, R. M. (2014). The effects of problem posing on student mathematical learning: A meta-analysis. *International Education Studies*, 7(13), 227
- Rowland, T., Turner, F., Thwaites, A. & Huckstep, P. (2009) *Developing Primary Mathematics Teaching: Reflecting on Practice with the Knowledge Quartet*. London: Sage.
- Ruthven, K. (1998). The Use of Mental, Written and Calculator Strategies of Numerical Computation by Upper Primary Pupils within a 'Calculator-Aware' Number Curriculum. *British Educational Research Journal*, 24(1), 21-42.
- Ruthven, K. (2001). British research on developing numeracy with technology. In M. Askew & M. Brown (Eds.), *Teaching and Learning Primary Numeracy: Policy, Practice and Effectiveness. A review of British research for the British Educational Research Association in conjunction with the British Society for Research into Learning of Mathematics* (pp. 29-33). Southwell, Notts: British Educational Research Association (BERA).

- Ruthven, K. (2009). Towards a calculator-aware mathematics curriculum. *Mediterranean Journal for Research in Mathematics Education*, 8(1), 111-124.
- Ruthven, K. (2015). Taking design to task: A critical appreciation. In A. Watson, & M. Ohtani (Eds.), *Task design in mathematics education: An ICMI study 22* (pp. 311-320). Heidelberg: Springer.
- Ryan, J., & Williams, J. (2007). *Children's mathematics 4-15: learning from errors and misconceptions*. Buckingham: Open University Press.
- * Sahin, B. (2016) Effect of the use of technology in mathematics course on attitude: A meta analysis study. *Turkish Online Journal of Educational Technology*, (November Special Issue), pp. 809-814.
- * Salinas, A. (2010). Investing in our teachers: What focus of professional development leads to the highest student gains in mathematics achievement? PhD Thesis, University of Miami
- * Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, 19, 158-172.
- * Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). Review and meta-analyses of school and teaching effectiveness. Enschede: Department of Educational Organisation and Management, University of Twente.
- * Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does Working Memory Training Transfer? A Meta-Analysis Including Training Conditions as Moderators. *Educational Psychologist*, 50(2), 138-166. doi: 10.1080/00461520.2015.1036274
- Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior modification*, 22(3), 221-242.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 465-494). New York: Macmillan.
- Shaughnessy, J. M. (2007). Research on Statistics Learning and Reasoning. In F. K. J. Lester (Ed.), *Second handbook of Research on mathematics teaching and learning* (pp. 957-1010). Greenwich, CT: Information Age Publishing.
- Shayer, M., & Adhami, M. (2007). Fostering Cognitive Development Through the Context of Mathematics: Results of the CAME Project. *Educational Studies in Mathematics*, 64(3), 265-291.
- Shayer, M., & Adhami, M. (2010). Realizing the cognitive potential of children 5–7 with a mathematics focus: Post-test and long-term effects of a 2-year intervention. *British Journal of Educational Psychology*, 80(3), 363-379. doi:10.1348/000709909x482363
- Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., . . . Boers, M. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology*, 62(10), 1013-1020. doi:10.1016/j.jclinepi.2008.10.009
- Shuard, H., Walsh, A., Goodwin, J., & Worcester, V. (1991). *Calculators, children and mathematics: The Calculator-Aware Number curriculum*. Hemel Hempstead: Simon & Schuster.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

- Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J. (2010). Developing effective fractions instruction for kindergarten through 8th grade: A practice guide (NCEE #2010-4039). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from whatworks.ed.gov/publications/practiceguides.
- Simms, V. (2016). Mathematical mindsets: unleashing students' potential through creative math, inspiring messages and innovative teaching. *Research in Mathematics Education*, 18(3), 317-320. doi:10.1080/14794802.2016.1237374
- Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26, 114
- Simpson, A. (2017). The misdirection of public policy: comparing and combining standardised effect sizes. *Journal of Education Policy*, 1-17. doi:10.1080/02680939.2017.1280183
- Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, 50, 315-342.
- Slavin, R. E. (1984). Team-Assisted individualization; Cooperative learning and individualized instruction in the mainstreamed classroom. *Remedial and Special Education*, 5, 33-42.
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57, 293-336.
- *Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, 57(2), 175-213. doi:10.3102/00346543057002175
- Slavin, R. E. (1993). Ability grouping in the middle grades: Achievement effects and alternatives. *The Elementary School Journal*, 93(5), 535-552.
- Slavin, R. E. (1999). Comprehensive approaches to cooperative learning. *Theory into Practice*, 38(2), 74-79.
- *Slavin, R. E. and Lake, C. (2007). *Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis* (Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University, 2007).
- †Slavin, R. E., & Lake, C. (2008). *Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis*. *Review of Educational Research*, 78(3), 427-515. doi:10.3102/0034654308317473.
- †Slavin, R. E., Groff, C., & Lake, C. (2009). *Effective Programs in Middle and High School Mathematics: A Best-Evidence Synthesis*. *Review of Educational Research*, 79(2), 839-911.
- *Slavin, R. E., Lake, C., & Groff, C. (2007). *A Best-Evidence Synthesis Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis*. Baltimore: The Best Evidence Encyclopedia, Center for Data-Driven Reform in Education, Johns Hopkins University.
- *Slavin, R. E., Lake, C., & Groff, C. (2009). *Effective programs in middle and high school mathematics: A best-evidence synthesis*. *Review of Educational Research*, 79(2), 839-911.
- Slavin, R. E., Sheard, M., Hanley, P., Elliott, L., & Chambers, B. (2013). *Effects of Co-operative Learning and Embedded Multimedia on Mathematics Learning in Key Stage 2: Final Report*. York: Institute for Effective Education.

- Smith III, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115-163.
- * Smith, B.A. (1996). A meta-analysis of outcomes from the use of calculators in mathematics education. (Doctoral dissertation, Texas A & M University-Commerce, 1996). *Dissertation Abstracts International*, 58(03), 787.
- Soderstrom, N. C., & Bjork, R. A. (2013). Learning versus performance. In D. S. Dunn (Ed.), *Oxford bibliographies online: Psychology*. New York, NY: Oxford University Press.
- * Sokolowski, A. (2015). The Effects of Mathematical Modelling on Students' Achievement-Meta-Analysis of Research. *IAFOR Journal of Education*, 3(1), 93-114.
- * Sokolowski, A., Li, Y., & Willson, V. (2015). The effects of using exploratory computerized environments in grades 1 to 8 mathematics: a meta-analysis of research. *International Journal of STEM Education*, 2(1), 1-17.
- Sowder, J. (1992). Estimation and number sense. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371-389). New York: Macmillan.
- * Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, 20, 498-505.
- Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: the mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, 31(2), 143-175.
- Star, J. R. (2005). Research Commentary: Reconceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education*, 36(5), 404-411.
- Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., Lyskawa, J., McCallum, W. G., Porath, J., & Zbiek, R. M. (2015). Teaching strategies for improving algebra knowledge in middle and high school students (NCEE 2014-4333). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: <http://whatworks.ed.gov>
- * Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970-987. doi:10.1037/a0032447
- † Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K–12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses. *Review of Educational Research*, 86(4), 849-899.
- Stein, M. (2001). Teaching and learning mathematics: How instruction can foster the knowing and understanding of number. In J. Brophy (Ed.), *Subject-specific instructional methods and activities* (Vol. 8, pp. 114-144). Amsterdam: JAI.
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp 319-369). Charlotte, NC: Information Age Publishing).

- Stein, M., Engle, R., Smith, M., & Hughes, E. (2008). Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell. *Mathematical Thinking And Learning*, 10(4), 313-340.
- Steinbring, H. (2011). Changed views on mathematical knowledge in the course of didactical theory development: independent corpus of scientific knowledge or result of social constructions? In T. Rowland & K. Ruthven (Eds.), *Mathematical Knowledge in Teaching* (pp. 43-64). Dordrecht, NL: Springer.
- * Stoner, D. A. (2004). The effects of cooperative learning strategies on mathematics achievement among middlegrades students: a meta-analysis. University of Georgia, Athens.
- Streefland, L. (1991). *Fractions in realistic mathematics education: A paradigm of developmental research* (Vol. 8). Dordrecht. The Netherlands: Kluwer Academic Publishers.
- Sullivan, P., Clarke, D., & Clarke, B. (2013). *Teaching with tasks for effective mathematics learning*. New York: Springer.
- Swan, M. (2005). *Improving learning in mathematics*. London: Department for Education and Skills.
- Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices*. London: National Institute for Advanced and Continuing Education (NIACE).
- Swan, M. (2008). A Designer Speaks. *Educational Designer*, 1(1). Retrieved from http://www.educationaldesigner.org/ed/volume1/issue1/article3/#s_2 on 11 May 2017.
- * Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject design intervention research for students with LD. *Journal of Learning Disabilities*, 33, 114–136.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4), 295-312.
- Symonds, J. E., & Galton, M. (2014). Moving to the next school at age 10–14 years: an international review of psychological development at school transition. *Review of Education*, 2(1), 1-27. doi: 10.1002/rev3.3021
- Takahashi, A. (2016). Recent Trends in Japanese Mathematics Textbooks for Elementary Grades: Supporting Teachers to Teach Mathematics through Problem Solving. *Universal Journal of Educational Research*, 4(2), 313-319.
- * Templeton, T. N., Neel, R. S., & Blood, E. (2008). Meta-analysis of math interventions for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 16(4), 226-239. <http://dx.doi.org/10.1177/1063426608321691>
- Thompson, I. (2001). British research on mental and written calculation methods for addition and subtraction. In M. Askew & M. Brown (Eds.), *Teaching and learning primary numeracy: Policy, practice and effectiveness. A review of British research for the British Educational Research Association in conjunction with the British Society for Research into Learning of Mathematics* (pp. 15-21). Southwell, Notts: British Educational Research Association (BERA).
- Thompson, S., & Senk, S. (2001). The effects of curriculum on achievement in second-year algebra: The example of the University of Chicago School

- Mathematics Project. *Journal for Research in Mathematics Education*, 32(1), 58-84.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). Teacher professional learning and development. Best evidence synthesis iteration (BES). Wellington, New Zealand: Ministry of Education.
- *Tingir, S., Cavlazoglu, B., Caliskan, O., Koklu, O., Intepe-Tingir, S. (In press) Effects of mobile devices on K-12 students' achievement: A meta-analysis. *Journal of Computer Assisted Learning*, DOI: 10.1111/jcal.12184
- Tobin, K. (1986). Effects of Teacher Wait Time on Discourse Characteristics in Mathematics and Language Arts Classes. *American Educational Research Journal*, 23(2), 191-200.
- Tobin, K. (1987). The Role of Wait Time in Higher Cognitive Level Learning. *Review of Educational Research*, 57(1), 69-95. doi:10.3102/00346543057001069
- *Tokac, U., Novak, E. & Thompson, C. (2015). Effects of Game-Based Learning on Students' Mathematics Achievement: A Meta-Analysis. Representing Florida State University with a poster presentation at 2015 Statewide Graduate Student Research Symposium, University of Central Florida, Orlando, FL, April 24, 2015.
- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013). Elementary School Mathematics intervention report: Peer-Assisted Learning Strategies. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In *Encyclopedia of mathematics education* (pp. 521-525). Springer Netherlands.
- Verschaffel, L., Torbeyns, J., & De Smedt, B. (Forthcoming). Young children's early mathematical competencies: analysis and stimulation. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10)*. Dublin: European Society for Research in Mathematics Education.
- Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, 78(3), 516-551.
- Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: The role of learner generated examples*. Mahwah, NJ: Erlbaum.
- Watson, A., & Mason, J. (2007). Taken-as-shared: A review of common assumptions about mathematical tasks in teacher education. *Journal of Mathematics Teacher Education*, 10, 205-215.
- Watson, A., & Ohtani, M. (Eds.). (2015). *Task design in mathematics education: An ICMI study 22*. Heidelberg: Springer.
- Watson, A., Jones, K., & Pratt, D. (2013). *Key Ideas in Teaching Mathematics: Research-based guidance for ages 9-19*. Oxford: Oxford University Press.
- *Waxman, H. C., Wang, M. C., Anderson, K. A., & Walberg, H. J. (1985). Adaptive Education and Student Outcomes: A Quantitative Synthesis. *The Journal of Educational Research*, 78(4), 228-236. doi:10.1080/00220671.1985.10885607
- What Works Clearinghouse (2006). *Elementary school math*. Washington, DC: U.S. Department of Education.

- What Works Clearinghouse (2007). Middle school math. Washington, DC: U.S. Department of Education.
- Wheater, R., Ager, R., Burge, B., & Sizmur, J. (2014). Achievement of 15-Year-Olds in England: PISA 2012 National Report (OECD Programme for International Student Assessment) [Revised Version, April 2014]. London: Department for Education.
- William, D. (2017). Memories are made of this. TES (2nd June 2017).
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- *Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review*, 22(4), 393-409.
- Woodward, J., Beckmann, S., Driscoll, M., Franke, M. L., Herzig, P., Jitendra, A., Koedinger, K. R., Ogbuehi, P. (2012). Improving mathematical problem solving in grades 4 through 8: A practice guide (NCEE 2012-4055). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- WSIPP (2017) Washington State Institute for Public Policy Benefit-Cost Results: Tutoring By Peers Pre-K to 12 Education.
- *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(4), 207-225.
- *Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Zanobini, M., & Usai, M. C. (2002). Domain-specific self-concept and achievement motivation in the transition from primary to low middle school. *Educational Psychology*, 22(2), 203-217.
- *Zhang, D., & Xin, Y. P. (2012). A follow-up meta-analysis for word-problem-solving interventions for students with mathematics difficulties. *The Journal of educational research*, 105(5), 303-318.

14 Appendix: Technical

Correlation, r : Another measure of effect size is a correlation coefficient r (or its squared value, r^2). The measure r varies between -1 for perfect negative linear correlation and 1 for perfect positive linear correlation. A value of zero represents no linear correlation. It is important to note that correlation does not necessarily imply causation, and that r captures only linear correlation, and cannot be used to measure more complicated non-linear associations between variables.

Percentage of non-overlapping data (PND): PND is a less-frequently-used kind of effect size, used for single-subject experimental designs. PND scores above 90 may be considered to represent “very effective” treatments, scores from 70 to 90 represent “effective” treatments, scores from 50 to 70 “questionable” treatments and scores below 50 “ineffective” treatments (Scruggs & Mastropieri, 1998, p. 224). It is important to note that PND is a less robust measure of effect size than d or g .

Publication bias: One possible source of inflation of effect sizes is *publication bias*. Also known as the ‘file-drawer problem’, this refers to the possibility that studies that lead to smaller effect sizes are less likely to be published, so that the published literature becomes biased towards higher effect sizes. Methods have been devised that attempt to identify and correct for publication bias, but these are imperfect.

Heterogeneity: When multiple effect sizes are compared across studies, it is often helpful to calculate Cochran’s Q , which is a measure of heterogeneity. When Q is large, the effect sizes are quite different from one another, which could suggest that there is more than one underlying effect, and they should not be considered as all measuring the same thing. In such cases, we often look for other variables (moderators) which may be able to account for some of the variation in the effect sizes across the studies.

Reference

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior modification*, 22(3), 221-242.

15 Appendix: Literature Searches

The literature searches were conducted between January and March 2017 using the search terms, databases and search strings set out below. In addition, we also carried out hand searches of journals such as *Review of Educational Research*, *Education Research Review* and *Review of Education*, as well as searches of references lists from relevant literature.

Search Terms		
General	Specific	Literature type
mathematic* math* numeracy arithmetic education pedagogy intervention* strateg* teach* learn* instruction	manipulative concrete apparatus imagery visualization* Diagram* textbook* resource* statistic* quantitative literacy math* anxiety professional development mastery transition bridging transfer	a meta-analysis a meta-analytic meta-analysis meta-analytic quantitative synthesis best evidence synthesis systematic review research review research synthesis review of research
Databases searched		
ArticleFirst OCLC, British Education Index, Child Development & Adolescent Studies, ECO, Education Abstracts, EducatiOnline, ERIC, JSTOR, MathSciNet via EBSCOhost, PapersFirst OCLC, PsycARTICLES, ProQuest, PsycINFO, Teacher Reference Center Google / Google Scholar		
Search strings (full text / Title):		Initial Hits (in full text / only in title):

(math* OR numeracy OR arithmetic) AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	1744 / 59
(math* OR numeracy OR arithmetic) AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("a meta-analysis" OR "a meta-analytic")	1008 / 45
(manipulative* OR imagery OR "concrete apparatus" OR visualization*) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	24 / 1
(textbook OR resource) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	101 / 0
(statistic* OR "quantitative literacy") AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	487 / 0
"math* anxiety" AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	14 / 0
mastery AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	12 / 0
(transition OR bridging) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	24 / 0
(transition OR bridging) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")	40 / 0

<p>("professional development" OR "in-service training") AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")</p>	4 / 0
<p>diagram AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review")</p>	26 / 0 609 / 41 Extracted: 31
<p>transfer AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("meta-analysis" OR "meta-analytic" OR "quantitative synthesis" OR "best evidence synthesis" OR "systematic review") (math* OR numeracy OR arithmetic) AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")</p>	18 / 1 Extracted: 2
<p>(manipulative* OR imagery OR "concrete apparatus" OR visualization* OR diagram*) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")</p>	62 / 0 Extracted: 1
<p>(textbook OR resource) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")</p>	95 / 0 Extracted: 1
<p>(statistic* OR "quantitative literacy") AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")</p>	10 / 0 Extracted: 1
<p>"math* anxiety" AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")</p>	2 / 0 Extracted: 0
<p>mastery AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction)</p>	18 / 0 Extracted: 0
	21 / 0

AND ("research review" OR "research synthesis" OR "review of research")

Extracted: 0

(transition OR bridging OR transfer) AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")

("professional development" OR "in-service training") AND math* AND (education OR pedagogy OR intervention* OR strateg* OR teach* OR learn* OR instruction) AND ("research review" OR "research synthesis" OR "review of research")

Google Scholar Searches	Hits
allintitle: mathematics education OR educational "meta analysis"	32
allintitle: mathematics education OR educational "meta analytic"	3
allintitle: mathematics education OR educational "quantitative synthesis"	0
	3
allintitle: mathematics education OR educational "best evidence synthesis"	4
	14
allintitle: mathematics education OR educational "systematic review"	3
allintitle: mathematics school "meta analysis"	0
allintitle: mathematics school "meta analytic"	5
allintitle: mathematics school "quantitative synthesis"	2
allintitle: mathematics school "best evidence synthesis"	2
allintitle: mathematics school "systematic review"	0
allintitle: mathematics intervention "meta analysis"	0
allintitle: mathematics intervention "meta analytic"	0
allintitle: mathematics intervention "quantitative synthesis"	1
allintitle: mathematics intervention "best evidence synthesis"	14
allintitle: mathematics intervention "systematic review"	2
allintitle: mathematics teacher OR teaching "meta analysis"	0
allintitle: mathematics teacher OR teaching "meta analytic"	1
allintitle: mathematics teacher OR teaching "quantitative synthesis"	3
allintitle: mathematics teacher OR teaching "best evidence synthesis"	4
allintitle: mathematics teacher OR teaching "systematic review"	0
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization "meta analysis"	0
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization "meta analytic"	0
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization "quantitative synthesis"	0
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization "best evidence synthesis"	0
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization "systematic review"	0
allintitle: mathematics mastery "meta analysis"	0
allintitle: mathematics mastery "meta analytic"	0

allintitle: mathematics mastery "quantitative synthesis"	0
allintitle: mathematics mastery "best evidence synthesis"	0
allintitle: mathematics mastery "systematic review"	0
allintitle: mathematics transition OR bridging "meta analysis"	0
allintitle: mathematics transition OR bridging "meta analytic"	0
allintitle: mathematics transition OR bridging "quantitative synthesis"	0
allintitle: mathematics transition OR bridging "best evidence synthesis"	3
allintitle: mathematics transition OR bridging "systematic review"	0
allintitle: transition OR bridging school OR education OR educational "meta analysis"	0
allintitle: transition OR bridging school OR education OR educational "meta analytic"	0
allintitle: transition OR bridging school OR education OR educational "quantitative synthesis"	0
allintitle: transition OR bridging school OR education OR educational "best evidence synthesis"	13
allintitle: transition OR bridging school OR education OR educational "systematic review"	0
allintitle: mathematics "professional development" "meta analysis"	0
allintitle: mathematics "professional development" "meta analytic"	0
allintitle: mathematics "professional development" "quantitative synthesis"	0
allintitle: mathematics "professional development" "best evidence synthesis"	0
allintitle: mathematics "professional development" "systematic review"	0
allintitle: mathematics diagram "meta analysis"	0
allintitle: mathematics diagram "meta analytic"	1
allintitle: mathematics diagram "quantitative synthesis"	0
allintitle: mathematics diagram "best evidence synthesis"	0
allintitle: mathematics diagram "systematic review"	0
allintitle: mathematics transfer "meta analysis"	0
allintitle: mathematics transfer "meta analytic"	5
allintitle: mathematics transfer "quantitative synthesis"	0
allintitle: mathematics transfer "best evidence synthesis"	0
allintitle: mathematics transfer "systematic review"	0

allintitle: transfer school OR education OR educational "meta analysis"	1
allintitle: transfer school OR education OR educational "meta analytic"	730
allintitle: transfer school OR education OR educational	40
"quantitative synthesis"	233
allintitle: transfer school OR education OR educational "best evidence synthesis"	731
allintitle: transfer school OR education OR educational "systematic review"	1390
allintitle: mathematics manipulatives OR imagery OR "concrete apparatus" OR visualization	39
allintitle: mathematics education imagery OR visualization	274
allintitle: mathematics mastery	
allintitle: mathematics transition OR bridging	
allintitle: mathematics "professional development"	
allintitle: mathematics diagram	
allintitle: mathematics transfer	
First 20 pages of titles and abstracts assessed by eye for each (Google Scholar):	448000
mathematics education "meta-analysis"	322000
mathematics school "meta-analysis"	147000
mathematics teaching "meta-analysis"	67000
mathematics education "meta-analytic"	54800
mathematics school "meta-analytic"	29900
mathematics teaching "meta-analytic"	44500
numeracy OR arithmetic education "meta-analysis"	43300
numeracy OR arithmetic school "meta-analysis"	28100
numeracy OR arithmetic teaching "meta-analysis"	11300
numeracy OR arithmetic education "meta-analytic"	11400
numeracy OR arithmetic school "meta-analytic"	7390
numeracy OR arithmetic teaching "meta-analytic"	

16 Appendix: Inclusion/Exclusion Criteria

Include if:

1. A: Maths/mathematics/numeracy in title / abstract and it is relevant to at least one RQ

OR

B: (Search further strategy): The topic is relevant to at least one RQ and we have only limited evidence in our dataset about the RQ: BUT only include if the paper then focuses sufficiently on mathematics (i.e., a significant number of mathematics studies and mathematics reported separately or as a moderator variable)

2. A: Focused on strategies/interventions in mathematics teaching & learning at KS2/KS3

OR

B: Strategies / interventions that are relevant to mathematics teaching & learning and have been sufficiently investigated in mathematics teaching & learning (so it would be a moderator variable or separately reported within the study or a significant number of studies are focused on mathematics).

3. Relevance to KS2/KS3 will be interpreted broadly. Studies should be relevant to topics taught at these ages in England. Unless there is a specific reason otherwise, we consider studies with students aged 5-16 to be relevant, although we would expect to express caution if only a limited number of studies were with the KS2/KS3 age group (i.e., ages 9-13).

Exclude if:

1. Not written in English
2. Meta-analysis published before 1970 (although original studies could be published before this date).
3. Concerned with students with specific learning difficulties (i.e., those lying outside the continuum of typical development).
4. Concerned with aspects of knowing / understanding / doing mathematics (or differences in gender, ethnicity, SES, etc.) but not about educational or teaching interventions that address these (although some of this literature might be relevant to the typical development section).
5. The paper cannot be located.

Examples of excluded meta-analyses:

Browder, D. M., Spooner, F., Ahlgrim-DeLzell, L., Harris, A. A., & Wakemanxya, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional children*, 74(4), 407-432. [Excluded due to focus on students with significant cognitive difficulties.]

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, 148, 163-172. [Excluded because not a meta-analysis of interventions.]

Hyde, Janet S.; Fennema, Elizabeth; Lamon, Susan J. (1990) Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, Vol 107(2), Mar 1990, 139-155. <http://dx.doi.org/10.1037/0033-2909.107.2.139> [[Excluded because not a meta-analysis of interventions.]

Kulik, C. L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of educational research*, 60(2), 265-299. [Excluded because not focused on mathematics.]

Ma, X., & Kishor, N. (1997). Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review. *Educational Psychology Review*, 9(2), 89-120. [Excluded because not a meta-analysis of interventions]

Torgerson, C. J., Porthouse, J., & Brooks, G. (2003). A systematic review and meta-analysis of randomised controlled trials evaluating interventions in adult literacy and numeracy. *Journal of Research in Reading*, 26(3), 234-255. Excluded because not relevant to Key Stage 2 and 3 (age 9-13.)

Contact information

Education Endowment Foundation (EEF)

9th Floor, Millbank Tower

London

SW1P 4QP

Email: info@eefoundation.org.uk

Nuffield Foundation

28 Bedford Square

London

WC1B 3JS

Email: info@nuffieldfoundation.org