Mathematics Teacher Exchange: China-England. Technical report and supplementary data and analysis

January 2019

Sheffield Hallam University

Boylan, M., Wolstenholme, C., Demack, S., Maxwell, B., Jay, T., Adams, G. and Reaney, S.

Glossary

Bivariate - a statistical term that refers to analyses that involve only two variables.

Cluster - a statistical term that refers to the grouping of similar units of analyses together. For example, pupils are clustered into schools and schools are clustered into geographical areas. This is an example of natural (systemic) hierarchical pupil clustering.

CPD - continuing professional development.

High mastery - a category derived from application of implementation criteria related to a model of MTE mastery pedagogy indicating implementation at a high level.

Inspire Maths - primary mathematics programme using translations of Singapore textbooks as core texts.

Linear Regression - a statistical modelling technique for examining variation in a scale/continuous outcome variable through the introduction of (scale or categorical) explanatory variables.

Logistic Regression - a statistical modelling technique for examining variation in a categorical outcome variable through the introduction of (scale or categorical) explanatory variables.

Mastery at or above threshold level - a category derived from application of implementation criteria related to a model of mastery pedagogy. When the meaning is clear in context this is shortened to implementation of mastery or similar.

Mastery specialist - an alumni of the Primary Mathematics Teaching for Mastery Specialists Programme with responsibility for leading change in their own school and supporting change in six to seven other schools, as well as collaborating with Maths Hub leadership and other mastery specialists.

Mathematics Mastery - primary mathematics programme, developed initially by the Ark Multi Academy Trust informed by Singapore mathematics curriculum and pedagogy.

Mathematics Teacher Exchange - exchange programme involving 48 English primary schools and teachers in Shanghai in 2014/15 and 70 English primary schools in 2016/17. Abbreviated as '**MTE**' or 'the exchange'.

Mathematics Teacher Exchange cohort 1 school - a school selected by the local Maths Hub which participated in the exchange in 2014/15 and hosted a Shanghai teacher and in nearly all cases had one or more members of staff visit Shanghai. In the first and third interim reports these schools were referred to as 'lead primary schools', however the change in terminology in the final report aims to avoid confusion with schools now identified in Maths Hubs as leading mastery developments, which in some cases are not the same. Abbreviated as '**MTE cohort 1 school**'.

Mathematics Teacher Exchange cohort 2 school - a school that was involved in the exchange in 2016/17 by having one of their members of staff visit Shanghai and, in most cases, hosting a Shanghai teacher. Teachers rather than schools were recruited to participate in the exchange programme and were selected from the alumni of the Primary Mathematics Teaching for Mastery Specialists Programme. Abbreviated as '**MTE cohort 2 school**'.

Mathematics Teacher Exchange lead - used to denote school staff who had been directly involved in the exchange programme and/or leading wider dissemination within their school and, in some cases, their local and wider Maths Hub Network. Note that in previous reports references were made to 'lead primary teacher'. However, as the Teaching For Mastery Programme has developed, leadership and promotion of teaching for mastery has extended to other teachers such as mastery specialists. Abbreviated as '**MTE lead**'.

Maths Hubs - a network of hubs across England each led or jointly led by a school or college. Maths Hubs work in partnership with neighbouring schools, colleges, universities, CPD providers, maths experts and employers. There were 32 Maths Hubs in England at the start of the exchange and as of November 2015 there are 35 Maths Hubs.

Maths No Problem - primary maths programme using translations of Singapore textbooks as core texts.

MTE Mastery pedagogy - the name given in the report to teaching approaches aiming to develop mastery informed by East Asian practices and used, in particular, to refer to practices of MTE schools. MTE mastery pedagogy is a more general description than the specific 'teaching for mastery' promoted by the NCETM.

Multilevel - a statistical term that relates to statistical modelling with more than one cluster level. A 2-level analysis might include school and individual pupil levels.

NCETM - National Centre for Excellence in the Teaching of Mathematics.

NCTL - National College for Teaching and Leadership.

Ofsted - Office for Standards in Education.

Primary Mathematics Teaching for Mastery Specialists Programme (PMTMSP) - intensive professional development programme for primary mathematics teachers led by the NCETM with 140 (with 133 completing) teachers participating in 2015/16, and 140 per year for six years from 2016/17.

SEND - Special Educational Need or Disability

Substantial mastery - self-reported implementation of mastery by interviewees. This is particularly important in analysis of impact where self-reports of substantial implementation of mastery for two years for the Y2 and Y6 2016/17 cohorts is used to define a sub-sample of schools for exploratory analysis.

Teaching for mastery - NCETM-promoted East Asian informed mastery pedagogy that is the focus of the PMTMSP. Abbreviated as **TfM**.

TSA - Teaching School Alliance - alliances led by a Teaching School, including schools benefiting from support and strategic partners. A **Teaching School** is an outstanding school that plays a leading role in the training and professional development of teachers, support staff and headteachers, as well as contributing to the raising of standards through school-to-school support.

References to previous evaluation reports:

The **'first interim report'** refers to the report of Boylan, Wolstenholme, Maxwell, Jay, Stevens and Demack (2016) Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Interim research report. (DfE)[1]

The **'second interim report'** refers to the report of Demack, Jay, Boylan, Wolstenholme, Stevens and Maxwell (2017) Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Second interim research report. (DfE)[2]

The **'third interim report'** refers to the report of Boylan, Maxwell, Wolstenholme and Jay (2017) Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Third interim research report. (DfE)[3]

[1] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/536003/Mathematics_Teacher_Exchange_Interim_Report_FINAL_040716.pdf

[2] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/666449/MTE_second_interim_report_121217.pdf

[3] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/666450/MTE_third_interim_report_121217.pdf

Contents

List of figures

List of tables

1. Introduction

This technical report is intended to be read alongside the final published main report of the Longitudinal evaluation of the Mathematics Teacher Exchange: China-England – Final Report and both together constitute a single research output. The table below provides a summary of the content of the sections that follow, and refers to the sections in the main report that material or data are most closely connected to.

Table 1: Content of technical report and corresponding area in main report

Technical report section	Content/purpose	Link to main report
Section 2. Previous reports	Briefly outlines the aims, objectives and contents of previous reports.	Section 1.4. Previous reports
Section 3. NCETM descriptions of mastery	NCETM descriptions of teaching for mastery are provided before the insertion of a copy of the NCETM's information for applicants for the Primary Mathematics Teaching for Mastery Specialists Programme.	Section 2 Mastery innovations and the Teaching For Mastery Programme and the Shanghai teacher exchange Section 4 The Mathematics Teacher Exchange as a change innovation
Section 4. Overview of the evaluation methodology	The four strands of the evaluation are summarised.	Section 3. Evaluation methodology, data collection and analysis
Section 5. Strand one Year 3 data collection and analysis and strand one data corpus	Details of the data collection and analysis for strand one are outlined, with more detail given on Year 3 qualitative data collection which has not been reported in previous interim evaluation reports.	Section 3. Evaluation methodology, data collection and analysis
Section 6. Strand one supporting data	Data predominantly collected through strand one is reported here, supporting findings discussed in the main report.	Section 6. Patterns of implementation and change over time
Section 7. Strand one	Discussion on how levels	Section 6. Patterns of

Technical report section	Content/purpose	Link to main report
implementation criteria and analysis	of implementation of mastery between schools have been determined.	implementation and change over time
Section 8. Strand two analysis	Effect sizes are reported and explained before further details of the impact analysis. Pupil attitude survey analysis is presented.	Section 8. Impact of change on pupils
Section 9. Strand four cohort 2 - evaluation data collection and analysis	Details of data collection and analysis for strand four: MTE cohort 2 schools.	Section 3.2. Collection and analysis of Year 3 data. Cohort 2 findings also reported throughout main report alongside cohort 1 findings.
Section 10. Further research into mastery implementation and the Teaching For Mastery Programme	Furher development of issues considered in the main report about addressing the need to gather further evidence.	Sectioin 13. Interpreation of findings.

2. Previous reports

Three interim reports have been published prior to the final and technical report. The content of these interim reports is outlined below.

First interim report[4]

The first interim report (published July 2016) presented predominantly qualitative data analysis of interviews with the 48 MTE cohort 1 school staff. Reporting focused on experiences of the exchange and schools' and teachers' initial implementation of Shanghai-informed mathematics teaching. The report:

- described and assessed the early impact of the first exchange on practices
- described and assessed the perceptions of pupil outcomes
- evaluated the efficacy of exchange activities
- gave an overview of survey data collected in 2014 from the MTE cohort 1 schools, as well as data from contrast and other Maths Hub schools.

The purpose of the survey was to identify and compare levels of mastery-informed teaching in schools both directly involved and not involved in the exchange. The first report also:

- gave an overview of the background of the MTE initiative and the aims of the exchange
- provided details on the longitudinal evaluation methodology
- identified issues to consider for the future success of the initiative.

Second interim report[5]

The second interim report (published December 2017) described the quasi-experimental design employed to examine attainment outcomes. Using propensity score matching a group of contrast school were identified to compare outcomes with those of the MTE exchange schools. The matching process is described. Limitations of the analytical

---

4 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/536003/Mathematics_Teacher_Exchange_Interim_Report_FINAL_040716.pdf

5 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/666449/MTE_second_interim_report_121217.pdf

approach are described in detail. The report then presents analysis of pupil outcome baseline data, through analysis of Key Stage 1 and Key Stage 2 attainment data. In order to form a baseline, data were utilised from a period prior to the start of the intervention until the end of the first year of the intervention. Data from the first Year 6 pupil attitude survey were also reported, providing a baseline to enable analysis of how the impact of any changes in practice influence changes in pupil attitudes to mathematics.

Third interim report[6]

The third interim report (published alongside the second report in December 2017) presented findings from analysis of follow-up interviews with lead teachers from the MTE cohort 1 schools. These took place during their second year of implementation in 2016. The report focused on changes in school-wide and classroom-based practice reported by teachers. The report also described variation of implementation across the schools, perceptions of teacher professional development outcomes and pupil outcomes.

6 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/666450/MTE_third_interim_report_121217.pdf

3. NCETM descriptions of mastery and teaching for mastery and the PMTMSP

For completeness and ease of reference, in this section the text of three NCETM descriptions of TfM are provided.

NCETM (2014). Mastery approaches to mathematics and the new national curriculum[7].

'Mastery' in high-performing countries

The content and principles underpinning the 2014 mathematics curriculum reflect those found in high-performing education systems internationally, particularly those of east and south-east Asian countries such as Singapore, Japan, South Korea and China. The OECD suggests that by age 15 students from these countries are, on average, up to three years ahead in maths compared to 15 year-olds in England.

What underpins this success is the far higher proportion of pupils reaching a high standard and the relatively small gaps in attainment between pupils in comparison to the picture in England.

Though there are many differences between the education systems of England and those of east and south-east Asia, we can learn from the mastery approach to teaching commonly followed in these countries. Certain principles and features characterise this approach:

- Teachers reinforce an expectation that all pupils are capable of achieving high standards in mathematics.
- The large majority of pupils progress through the curriculum content at the same pace. Differentiation is achieved by emphasising deep knowledge and through individual support and intervention.
- Teaching is underpinned by methodical curriculum design and supported by carefully crafted lessons and resources to foster deep conceptual and procedural knowledge.
- Practice and consolidation play a central role. Carefully designed variation within this builds fluency and understanding of underlying mathematical concepts in tandem.

[7] https://www.ncetm.org.uk/public/files/19990433/Developing_mastery_in_mathematics_october_2014.pdf.

- Teachers use precise questioning in class to test conceptual and procedural knowledge, and assess pupils regularly to identify those requiring intervention so that all pupils keep up.

The intention of these approaches is to provide all children with full access to the curriculum, enabling them to achieve confidence and competence – 'mastery' – in mathematics, rather than many failing to develop the maths skills they need for the future.

Curriculum changes

The 2014 national curriculum for mathematics has been designed to raise standards in maths, with the aim that the large majority of pupils will achieve mastery of the subject. Mathematics programmes of study state that:

- All pupils should become fluent in the fundamentals of mathematics, including through varied and frequent practice, so that pupils develop conceptual understanding and are able to recall and apply their knowledge rapidly and accurately to problems.
- The expectation is that the majority of pupils will move through the programmes of study at broadly the same pace. When to progress should always be based on the security of pupils' understanding and their readiness to progress to the next stage.
- Pupils who grasp concepts rapidly should be challenged through rich and sophisticated problems before any acceleration through new content. Those pupils who are not sufficiently fluent with earlier material should consolidate their understanding, including through additional practice, before moving on.

For many schools and teachers, the shift to this 'mastery curriculum' will be a significant one. It will require new approaches to lesson design, teaching, use of resources and support for pupils.

Key features of the mastery approach

Curriculum design

A detailed, structured curriculum is mapped out across all phases, ensuring continuity and supporting transition. Effective mastery curricula in mathematics are designed in relatively small, carefully sequenced steps, which must each be mastered before pupils move to the next stage. Fundamental skills and knowledge are secured first. This often entails focusing on curriculum content in considerable depth at early stages.

Teaching resources

A coherent programme of high quality curriculum materials is used to support classroom teaching. Concrete and pictorial representations of mathematics are chosen carefully to help build procedural and conceptual knowledge together. Exercises are structured with great care to build deep conceptual knowledge alongside developing procedural fluency.

The focus is on the development of deep structural knowledge and the ability to make connections. Making connections in mathematics deepens knowledge of concepts and procedures, ensures what is learnt is sustained over time, and cuts down the time required to assimilate and master later concepts and techniques.

One medium for coherent curriculum materials is high quality textbooks. These have the additional advantage that pupils also use them to return to topics studied, for consolidation and for revision. They represent an important link between school and home.

Lesson design

Lessons are crafted with similar care and are often perfected over time with input from other teachers, drawing on evidence from observations of pupils in class.

Lesson designs set out in detail well-tested methods to teach a given mathematical topic. They include a variety of representations needed to introduce and explore a concept effectively and also set out related teacher explanations and questions to pupils.

Teaching methods

Precise questioning during lessons ensures that pupils develop fluent technical proficiency and think deeply about the underpinning mathematical concepts. There is no prioritisation between technical proficiency and conceptual understanding; in successful classrooms these two key aspects of mathematical learning are developed in parallel.

Pupil support and differentiation

Taking a mastery approach, differentiation occurs in the support and intervention provided to different pupils, not in the topics taught, particularly at earlier stages. There is no differentiation in content taught, but the questioning and scaffolding individual pupils receive in class as they work through problems will differ, with higher attaining pupils challenged through more demanding problems which deepen their knowledge of the same content. Pupils' difficulties and misconceptions are identified through immediate formative assessment and addressed with rapid intervention – commonly through individual or small group support later the same day. There are very few 'closing the gap' strategies, because there are very few gaps to close.

Productivity and practice

Fluency comes from deep knowledge and practice. Pupils work hard and are productive. At early stages, explicit learning of multiplication tables is important in the journey towards fluency and contributes to quick and efficient mental calculation. Practice leads to other number facts becoming second nature. The ability to recall facts from long term memory and manipulate them to work out other facts is also important.

All tasks are chosen and sequenced carefully, offering appropriate variation in order to reveal the underlying mathematical structure to pupils. Both class work and homework provide this 'intelligent practice', which helps to develop deep and sustainable knowledge.

Implications for professional development and training of teachers

Teachers of mathematics in countries that perform well in international comparisons are mathematics specialists, including those in primary schools. They have deep subject knowledge, and deep knowledge of how to teach mathematics. They engage in collaborative planning and are continually seeking to improve their effectiveness.

Specialist mathematics teachers will therefore require:

- Deep structural subject knowledge of mathematics.
- Strong understanding of the structure of the curriculum and its aims: fluency, accuracy, precision, reasoning and problem-solving, and how to apply these to teaching.
- Insight of what is meant by mastery of the curriculum, the factors that contribute to it and how it is achieved.
- Appreciation of the importance of practice and the nature of intelligent practice to develop deep and sustainable understanding which contributes to mastery for all.
- Effective strategies to support pupils to learn, recall and apply multiplication tables.
- Knowledge of mathematics as a network of interconnected ideas and an appreciation that making connections reduces the amount of mathematics to learn, deepens knowledge and contributes to sustainability of understanding over time.
- The ability to select and employ effectively the use of mathematical representations to enable pupils to access the underlying structure of the mathematics.
- An appreciation of the features of good textbooks and when and how to use them appropriately to support high quality teaching.
- Opportunities to collaborate with other professionals.

- Knowledge of how effectively to deliver high quality whole-class teaching and provide access for all pupils.
- The ability to provide quick feedback to pupils and effective intervention to support all pupils to keep pace with the rest of the class.

NCETM (2016). The essence of mathematics teaching for mastery[8].

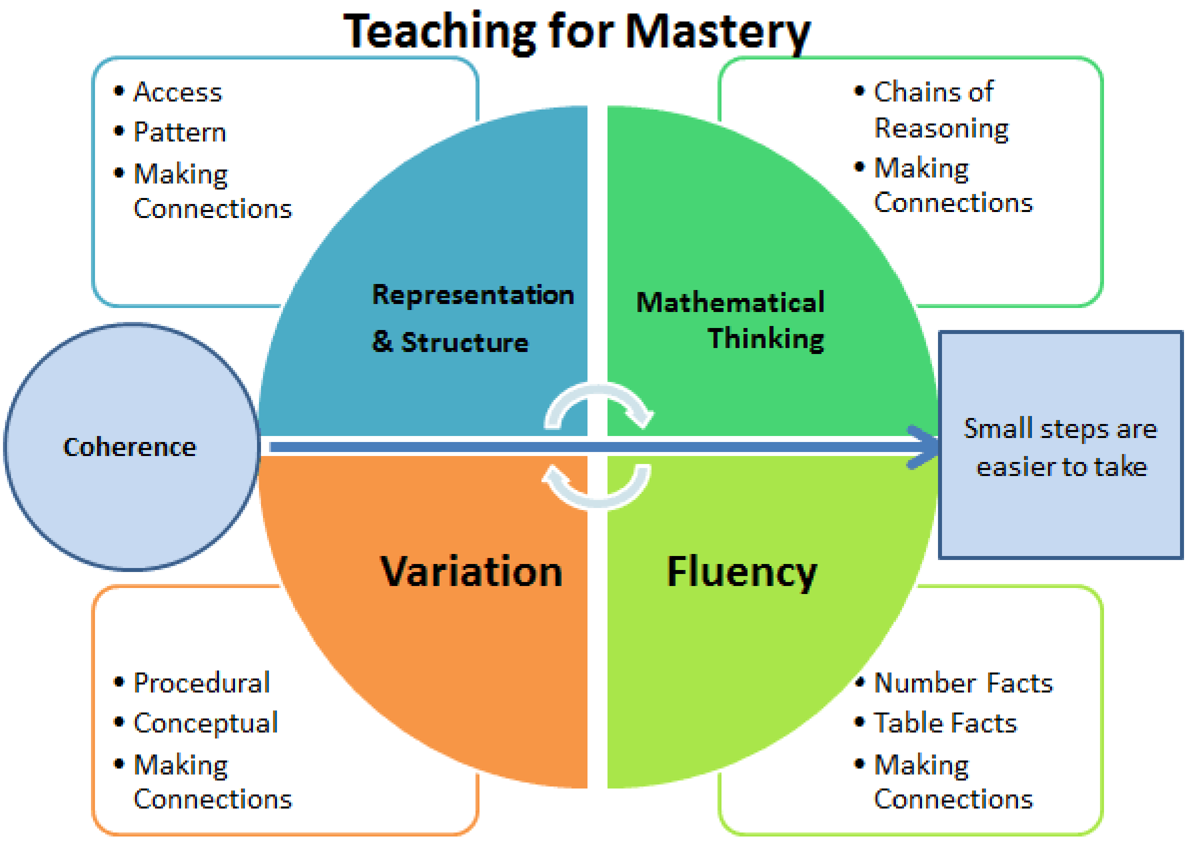
- Maths teaching for mastery rejects the idea that a large proportion of people ‘just can’t do maths’.
- All pupils are encouraged by the belief that by working hard at maths they can succeed.
- Pupils are taught through whole-class interactive teaching, where the focus is on **all** pupils working together on the same lesson content at the same time, as happens in Shanghai and several other regions that teach maths successfully. This ensures that all can master concepts before moving to the next part of the curriculum sequence, allowing no pupil to be left behind.
- If a pupil fails to grasp a concept or procedure, this is identified quickly and early intervention ensures the pupil is ready to move forward with the whole class in the next lesson.
- Lesson design identifies the new mathematics that is to be taught, the key points, the difficult points and a carefully sequenced journey through the learning. In a typical lesson pupils sit facing the teacher and the teacher leads back and forth interaction, including questioning, short tasks, explanation, demonstration, and discussion.
- Procedural fluency and conceptual understanding are developed in tandem because each supports the development of the other.
- It is recognised that practice is a vital part of learning, but the practice used is **intelligent practice** that both reinforces pupils’ procedural fluency and develops their conceptual understanding.
- Significant time is spent developing deep knowledge of the key ideas that are needed to underpin future learning. The structure and connections within the mathematics are emphasised, so that pupils develop deep learning that can be sustained.
- Key facts such as multiplication tables and addition facts within 10 are learnt to automaticity to avoid cognitive overload in the working memory and enable pupils to focus on new concepts.

---

[8] URL https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf Retrieved July 2016.

NCETM (2017). Five Big Ideas in Teaching for Mastery[9].

A central component in the NCETM/Maths Hubs programmes to develop Mastery Specialists has been discussion of Five Big Ideas, drawn from research evidence, underpinning teaching for mastery. This is the diagram used to help bind these ideas together:



A true understanding of these ideas will probably come about only after discussion with other teachers and by exploring how the ideas are reflected in day-to-day maths teaching, but here's a flavour of what lies behind them:

Coherence
Connecting new ideas to concepts that have already been understood, and ensuring that, once understood and mastered, new ideas are used again in next steps of learning, all steps being small steps

[9] URL https://www.ncetm.org.uk/resources/50042

Representation and Structure
Representations used in lessons expose the mathematical structure being taught, the aim being that students can do the maths without recourse to the representation

Mathematical Thinking
If taught ideas are to be understood deeply, they must not merely be passively received but must be worked on by the student: thought about, reasoned with and discussed with others

Fluency
Quick and efficient recall of facts and procedures and the flexibility to move between different contexts and representations of mathematics

Variation
Varying the way a concept is initially presented to students, by giving examples that display a concept as well as those that don't display it. Also, carefully varying practice questions so that mechanical repetition is avoided, and thinking is encouraged.

Primary Mastery Specialist Programme: Cohort 3 (2017-18) – Information

The text below is taken from the NCETM's information for applicants for the Primary Mastery Specialist Programme

Following the very successful first two cohorts of the Mastery Specialist programme, the NCETM and Maths Hubs are now seeking to recruit a third cohort of 140 expert primary school teachers (4 per Maths Hub) to develop and work as **Primary Mastery Specialists**. This document gives information about the programme and how teachers and their schools can apply to be involved. The closing date for applications is Wednesday 19th April.

Background

Since 2014, The NCETM and Maths Hubs have been working together to develop approaches to teaching for mastery within primary mathematics. This has been informed by the teaching of mathematics in high performing South East Asian jurisdictions. Each year since 2015 the NCETM and Maths Hubs have recruited 140 Mastery Specialist teachers. The first year in post is a training and development year, establishing mastery in their own schools, with the support of the senior leadership team. In the following years, they lead Work Groups involving six or seven other schools in the development of teaching for mastery.

The programme

The NCETM and the Maths Hubs recruited a third cohort of 140 teachers to develop and work as Mastery Specialists. 2017-18 was the development year for the teachers and their schools and then, in 2018-19 and in 2019-20, the Mastery Specialists will each lead a Teaching for Mastery Work Group for their Maths Hub. It is expected that the Mastery Specialist's school becomes a leading exponent of teaching for mastery in this time and so the school needs to ensure that it has the capacity and desire to take on and develop a teaching for mastery approach in the next few years. Before a teacher applies for the role, head teachers might find it useful to look at the NCETM website where there are interviews with heads who have led the introduction of teaching for mastery in a school. (https://www.ncetm.org.uk/resources/49822)

In their development year, 2017-18, participating teachers will:

- attend the NCETM cohort induction day (Monday 10th July 2017) along with their head teacher
- participate in three two-day residential professional development events led by the NCETM (October 2017, January 2018, and June 2018)
- develop their own understanding and skills for teaching mathematics for mastery in their own class
- work with colleagues, supported by the head teacher, to develop teaching for mastery approaches across their school, using a range of professional development activity, including regular Teacher Research Group (TRG) meetings
- lead a pilot TRG with teachers from interested local schools
- collaborate with the Maths Hub's leadership and the other Maths Hub Mastery Specialists

This will require 15 days teacher release time and will be fully funded through the Maths Hubs.

In 2018-19 and 2019-20, the Mastery Specialists will:

- lead a Teaching for Mastery Work Group for their Maths Hub involving six/seven schools
- lead a half-termly Work Group meeting with 12/14 teachers (two lead teachers from each school)
- carry out a termly support visit to each school to observe teaching, support in-school TRGs, and work with the lead teachers and head teacher

- continue to develop and embed teaching for mastery approaches across their own school
- continue to collaborate with the Maths Hub's leadership and Mastery Specialists

This will require 30/33 days teacher release time and will be fully funded through the Maths Hubs.

Benefits for participating teachers and their schools

Participating in the programme will provide the following benefits to the Mastery Specialists and their schools:

- Mastery Specialists will develop:
    - understanding of the principles of mastery within the context of teaching mathematics
    - deep subject knowledge of primary mathematics to support teaching for mastery
    - skills of teaching, planning and assessment in order to effectively support pupils in developing a deep and sustainable understanding (i.e. mastery) of mathematics
    - ability to support teachers, within their own school and in other schools, to adopt a teaching for mastery approach, including leading Teacher Research Groups
- Mastery Specialists will have the opportunity to work closely with the NCETM team and the national and local communities of Mastery Specialists
- Mastery Specialists, who are not already accredited NCETM PD Leads, will be able to gain this accreditation through successful completion of the programme
- The Mastery Specialist's school will benefit from high quality and sustained support in embedding teaching for mastery across the school

Who should apply

Table 2 below shows the essential and desirable criteria for applicants to the programme. This should be evidenced in the application form, which includes both the applicant's statement and the head teacher's reference.

Table 2: Criteria for applicants to the mastery specialist programme

Essential	Desirable
Qualified Teacher Status	Additional Status, e.g. Mathematics SLE/MaST
Employed as a teacher in a	Mathematics Subject Leader

Essential	Desirable
Primary/Infant/Junior/Middle School, and regularly teaching mathematics to the same class of children at least three days/week	
Able to fulfil the programme requirements and time commitment outlined above	
Good teaching skills in mathematics as evidenced by internal/external/Ofsted observation	A minimum of four years teaching primary mathematics
Passion and enthusiasm for teaching for mastery	A desire to develop as a specialist teacher of primary mathematics
Ability to work collaboratively with others	
Successful track record of working with other professionals effectively within your own school	Successful track record of working effectively with other professionals across a group of schools
Excellent communication and interpersonal skills	The ability to grow leadership capacity in others
An understanding of what constitutes effective learning in mathematics and the ability and confidence to communicate this	

Maths Hubs will look to appoint Mastery Specialists so that they both meet the criteria as set out above but also fit into the Hub's strategic plan for developing teaching for mastery across the region.

Expectations of participating teachers and their schools

For teachers selected to be part of the programme, there are the following expectations of them and their schools:

- The head teacher and Mastery Specialist commit to developing and embedding mathematics teaching for mastery approaches across the school, supported by professional development activity, including regular TRG meetings in their own school
- The head teacher commits to supporting the Mastery Specialist with their outreach work with other schools in ensuring that they are given appropriate release time .
- The Mastery Specialist commits to developing their understanding and practice related to mathematics teaching for mastery including:
    - attending the cohort induction day (10th July 2017)
    - participating in the three two-day residentials

  - developing mathematics teaching for mastery within their own class
- The Mastery Specialist commits to all aspects of the role and the release time required (2017-18: 15 days; 2018-19: 30/33 days; 2019-20: 30/33 days) including:
  - supporting teachers within their own school and leading regular TRG meetings
  - running a pilot TRG with interested schools (2017-18)
  - leading a Teaching for Mastery Work Group (2018-19 and 2019-20) involving half-termly cross-school TRG meetings and termly support visits to Work Group schools
  - collaborating with the Maths Hub's leadership and Mastery Specialists
- The head teacher commits to support the Mastery Specialist, including:
  - attending the cohort induction day (10th July 2017)
  - helping the specialist develop and embed teaching for mastery within the school
  - ensuring the teacher receives the required release time
- The Mastery Specialist and head teacher will provide any reports required by the Maths Hub and participate in any evaluation processes required

Funding

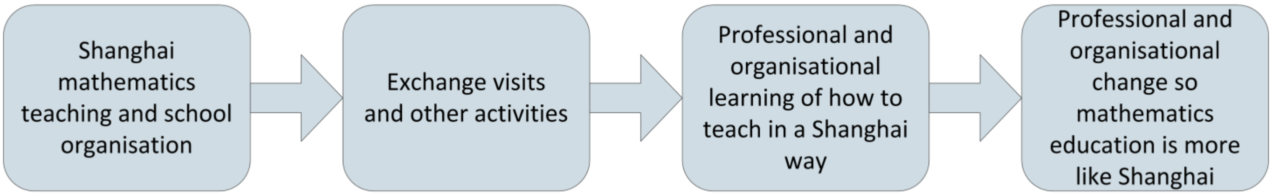
In the development year, 2017-18, your Maths Hub will fund the cost of 15 days release time for the Mastery Specialist's work and the cost of travel to the NCETM induction conference. There will also be £2000 matched funding for the school to purchase textbooks from the DfE approved list. In 2018-19 and in 2019-20, your Maths Hub will fund the cost of 30/33 days release time for the Mastery Specialist's work and the cost of travel for school support visits.

4. Overview of the evaluation methodology

4.1 The type of innovation

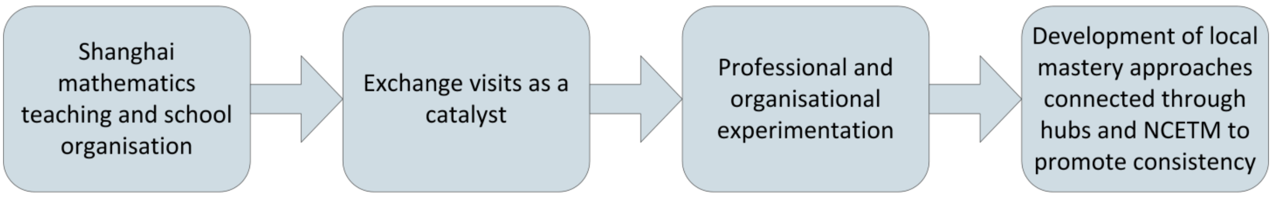
The MTE can be conceptualised in two different ways in terms of the nature of innovation and both of these perspectives inform the evaluation design. Firstly, the exchange has features of the implementation of a relatively well-defined innovation. From this perspective, the aim of the exchange is to *adopt* aspects of the Shanghai teaching approach. This is described in the figure below.

Figure 1: The Mathematics Teacher Exchange as an adoptive innovation



Alternatively, the exchange can be viewed as aiming to *adapt and develop* aspects of Shanghai mathematics education. This is shown in the figure below.

Figure 2: The Mathematics Teacher Exchange as an adaptive innovation



From this perspective, the primary aim of exchange visits and other activities is not to lead to professional skills and knowledge of how to teach or organise learning in a Shanghai way. Rather, it is to provide a stimulus to catalyse change that leads to professional and organisational learning through adaptation. Analysis of the exchange design, implementation and outcomes indicate that the exchange has features of both adoptive and adaptive innovation. Actual implementation of lessons from the exchange has been influenced by other factors including, importantly, NCETM's formulation and promotion of TfM.

4.2 Evaluation strands

The evaluation methodology is described in the main report, as are the evaluation objectives, and in more detail in the previous interim reports. In summary, the evaluation had four strands that are summarised here.

Strand one

Strand one consisted of a longitudinal multiple-case study design focused on MTE cohort 1 schools, encompassing both exploratory and evaluative dimensions (Yin, 2013). Data were collected through a combination of site visits and telephone interviews in three periods in the spring/summer of 2015, 2016 and 2017. This was supplemented, in the first year of the evaluation, by a set of interviews with Maths Hub leads and key NCETM and DfE stakeholders. In 2015 and 2017, mathematics coordinators in MTE cohort 1 schools and others within the Maths Hubs were surveyed. For 28 schools, reports were received either direct from schools or from NCETM in 2015. The NCETM also provided an analysis and summary of all end-of-year reports received by them as well as of schools' interim reports. Data from an NCETM survey in 2016 were also analysed.

Strand two

Strand two consisted of a longitudinal analysis of Key Stage 1 and Key Stage 2 attainment data, in comparison with a sample of 940 contrast schools[10]. Data used for the impact analysis were retrieved from the National Pupil Database (NPD) and the school census database. In addition, pupil attitudes to mathematics and to mathematics learning were surveyed in a sample of MTE cohort 1 schools in 2015 and 2017. The aim was to assess how changes to practices affect pupils' attitudes to mathematics over time. More detail is given in the main report in section 3.2.

Strand three

Strand three sought to identify initial patterns of effective change and early evidence of pupil impact in the first year of implementation. This involved follow-up telephone interviews with exchange teachers in a purposeful sample of MTE cohort 1 schools. Schools selected were ones where the initial case study visit indicated that notable changes in practices were occurring as a result of the exchange and this assessment could potentially be supported by school data. A thematic case analysis was conducted for each of the five MTE cohort 1 schools drawing on data collected during the strand one case study visit, together with data from the follow-up telephone interview. The five strand three cases also informed the analysis of patterns of implementation. Outcomes of strand three were reported in the first interim report.

[10] See below 'Analysis, reporting and samples' for further detail on the contrast school sample.

Strand four

Following the decision to extend the MTE to further cohorts and to embed it in the PMTMSP, the DfE commissioned an extension to the evaluation to consider the experiences and activities of a sample of MTE cohort 2 schools and teachers.

5. Strand one Year 3 data collection and analysis and strand one data corpus

In section 8, strand two data analysis is outlined. Below is the data collection and analysis for strand one.

5.1 Strand one Year 3 data collection and analysis

The final year of data collection in the MTE cohort 1 schools took place in spring/summer 2017 when schools were in their third year of implementation. A telephone interview was conducted with a member of staff from 40 of the 48 schools. The remaining eight schools withdrew from the research or did not respond to an invitation to be interviewed. The average length of the interviews was 59 minutes.

The 2017 interview schedule was made up of a mixture of 'checking' questions, closed questions and more in-depth, open questions. The checking questions were to confirm participants' answers from the 2016 interviews and to check if any particular practices had changed since then. For example, when asking about intervention, the interviewee would be read out the approaches to intervention they had described the previous year such as: intervention taking place daily, pupils being identified for intervention on a daily basis, lessons not being split, and intervention taking place after the lesson. Interviewees could then simply confirm this was the same or declare any changes to practice. This approach enabled a greater degree of confidence in the data, given that interviewees may have been substituted by colleagues over the course of the three-year project.

Closed questions were asked in order to help to quantify the changes made across the schools, and the open questions were used to gain in-depth information about the types of changes and the rationales for making them, and to obtain data in areas asked about in less detail in previous interviews such as changes to lesson preparation.

As was the procedure for the 2016 qualitative fieldwork, immediately following interviews, fieldworkers entered interviewees' answers to closed and checking questions into a spreadsheet. The spreadsheet data were then checked by the evaluation project manager to ensure these correlated with the transcripts. Where there was a discrepancy in the answers, a project director was asked to make the final decision. This process was utilised for ‘new’ closed questions as opposed to the checking questions as these were felt to be secure after the previous year’s checks.

All interviews recorded were then fully transcribed and the transcripts uploaded onto Nvivo 10 for analysis. An analysis meeting with the full team took place after the completion of all data collection, which enabled in-depth discussion about the implementation trends across the interviews, and aided in the development of a conceptual framework. This early meeting helped to develop an emerging understanding of implementation of practices across the schools.

Within Nvivo, it was important to test if the codes already developed for the previous stage of data collection and analysis were relevant and sufficient for the 2017 data collection. To do this, each member of the evaluation team took an interview transcript and attempted to code the contents to existing higher-level codes. In addition, each member looked in depth at a specific area of implementation to decide if existing codes and child codes were sufficient or if new ones needed to be created. The evaluation team (consisting of project directors and the project manager) held a further analysis meeting to discuss the outcomes of this activity and the need for new codes in particular areas. New codes were created where needed. For example in 2017, participants were asked about lesson preparation in much greater depth than in previous interviews and therefore additional codes were created.

5.2 Strand one data corpus

MTE cohort 1 schools' participation in interviews over the three years is summarised in Table 3 below. As can be seen, a total of 38 (out of 48) interviewees have taken part in an interview in all three years of the longitudinal evaluation. Teachers from two schools re-engaged with the evaluation in year 3 after declining an interview in year 2. Reasons for non-participation in interviews were related to staff availability and workload.

MTE cohort 1 participation in longitudinal evaluation

Table 3: Participation in interviews over 3 years of evaluation

All 3 years	Years 1 and 2	Years 1 and 3	Year 1 only
38	5	2	3
79%	11%	4%	6%

Source: MTE cohort 1, interviews 2015, 2016 and 2017

Analysis, reporting and samples

In 2014/15, 48 schools from 32 Maths Hubs participated in the first MTE with a total of 64 teachers and school leaders visiting Shanghai alongside additional educators and NCETM delegates. All of these schools participated in interviews in the first year (2015). In 2016 and 2017, not all of these schools responded to invitations to participate in interviews. Figure 3 below shows the numbers participating in one or other year and in both years.

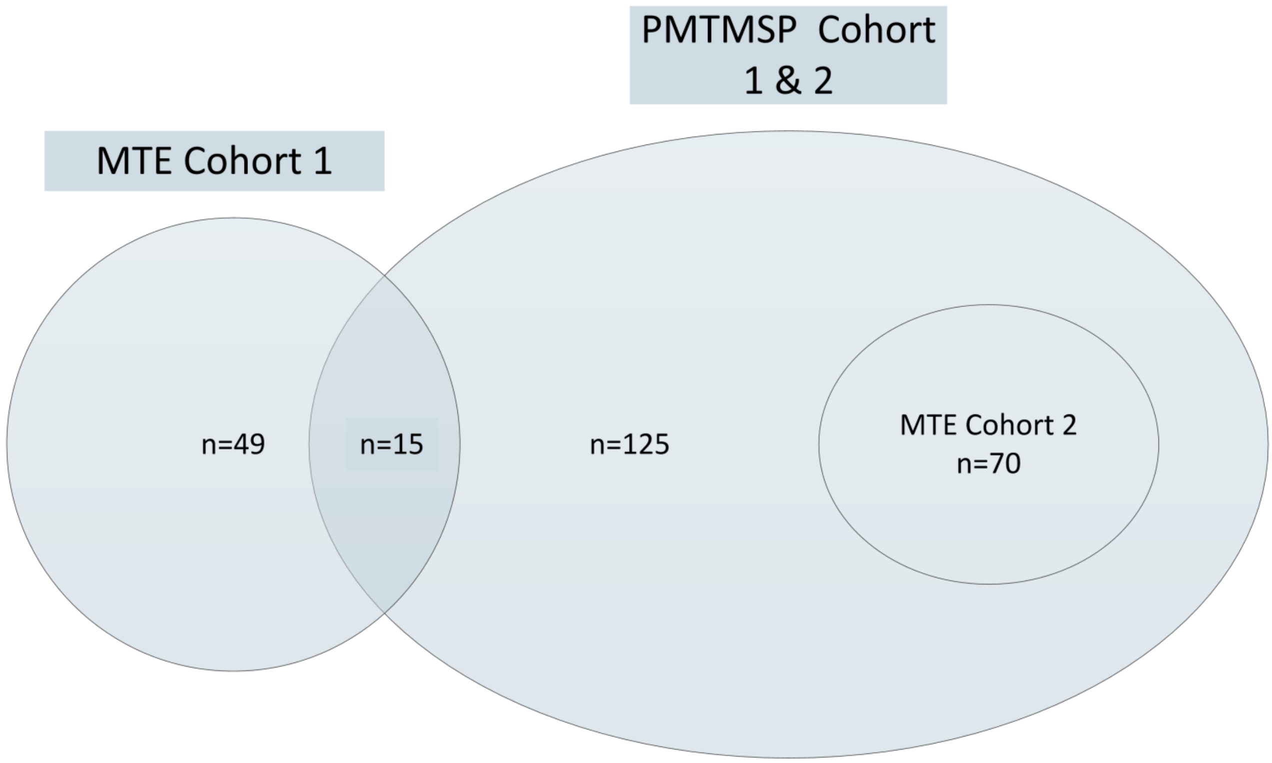
Figure 3: School participation in MTE cohort 1 interviews in 2016 and 2017

<table>
<tr><td></td><td></td><td colspan="3">2017</td></tr>
<tr><td></td><td></td><td>No</td><td>Yes</td><td>Total</td></tr>
<tr><td rowspan="3">2016</td><td>No</td><td>3</td><td>2</td><td>5</td></tr>
<tr><td>Yes</td><td>5</td><td>38</td><td>43</td></tr>
<tr><td>Total</td><td>8</td><td>40</td><td></td></tr>
</table>

Data from 2016 and 2017 were used in the analysis of levels of implementation of some aspects of practice and the overall determination as to whether or not MTE mastery pedagogy had been put in place. Consequently, in relation to implementation, data are reported for 38 schools in the main report. The technical report provides data where relevant on the larger samples. In reporting other issues such as constraining influences and supporting influences on implementation, all data (where relevant) is reported.

As discussed in the second interim report, although 48 schools participated in the MTE in 2014/15, one of these was an infant school and another was a co-located junior school. As reported in the second interim report, this was the only infant school in the evaluation. The propensity score matching was undertaken using 2014 school-level Key Stage 2 data and so this infant school was not included in this matching process (and hence any of the impact analyses). The junior school was included along with two other junior schools and 44 primary schools. As a consequence, the impact analyses included 47 MTE schools (44 primaries, three juniors). Figure 4 below provides details of the overall samples for number of teachers involved. Teachers from the sample of 48 MTE cohort 1 schools which participated in the PMTMSP may have been different to those who went on the exchange to Shanghai.

Figure 4: Number of teachers in the MTE cohort 1 and cohort 2 sample and PMTMSP



The figure described - for the PMTSMP recruited totals not those completing the programme which may be slightly lower. Details of other samples, such as schools that returned pupil attitude survey data, and other survey samples are reported in the relevant sections with further detail in the technical report.

In addition to data collection via interviews, 28 school reports were received either direct from schools or from NCETM in August 2015 in the first year of the evaluation. The NCETM also provided an analysis and summary of all end-of-year reports received by them as well as of schools' interim reports. In 2016 (the second year of the evaluation) these data were collected by an NCETM survey and passed to the evaluators by the NCETM. There were a total of 39 responses from schools; these data were used to triangulate findings form the second round of data collection. In 2017 no school report data were collected.

Other data collection methods for strand one, as described earlier in section 4 were:

- **Maths Hub lead interviews**: In year 1, telephone interviews with 12 Maths Hub leads took place between February and March 2015. Hubs were randomly selected and then contacted to invite them to interview. A thematic analysis of the Maths Hub lead interviews was undertaken.

- **Key stakeholder interviews**: In year 1, between February and March 2015, four interviews with key stakeholders from the NCETM and DfE were undertaken. Interviews were analysed thematically.
- **Mathematics coordinator survey**: A survey of all MTE cohort 1 schools, the 940 contrast schools, and other schools within Maths Hubs (identified by Maths Hub leads as having had some contact with Maths Hub activity and/or with the MTE schools) took place between June and July 2015 in year 1 of the evaluation. A total of 46 maths coordinators in the 48 MTE cohort 1 schools completed the first survey along with 218 Maths Hub schools, and 53 contrast schools. In order to identify change over time, the survey was repeated in June and July 2017 and was distributed to all MTE cohort 1 and 2 schools, 940 contrast schools and other schools within Maths Hubs (identified by Hub leads as having had some contact with Maths Hub activity and/or with the MTE schools). A total of 77 schools from cohort 1 and cohort 2 completed the first survey along with 33 contrast schools.

Table 4 below provides details of the data corpus for strand one of the evaluation for the MTE cohort 1 schools and pupil survey data collected for strand two.

Table 4: Data corpus MTE cohort 1 schools for strand one and pupil survey in strand two

Case code	2015 int (visit/tele int)	2016 int	2017 int	Coordinator survey 1	Coordinator survey 2	NCETM report Y1	NCETM report Y2	Pupil survey Year 1	Pupil survey year2
1a	V	Y	Y	Y	N	Y	Y	Y	N
1b	T	Y	Y	Y	Y	Y	Y	Y	Y
2a	V	Y	N	Y	Y	Y	Y	Y	N
2b	T	Y	Y	Y	N	Y	N	N	Y
3a	T	Y	Y	Y	N	Y	N	Y	Y
3b	T	Y	Y	Y	N	Y	Y	N	N
4a	V	Y	Y	Y	Y	Y	Y	Y	Y
5a	V	Y	Y	Y	Y	Y	Y	Y	N
6a	V	Y	Y	Y	Y	Y	Y	Y	N
7a	V	N	Y	Y	Y	Y	Y	Y	Y
8a	V	Y	Y	Y	N	Y	N	N	N
9a	V	Y	Y	Y	Y	Y	Y	N	N
9b	V	Y	Y	Y	N	Y	N	N	N
10a	V	Y	Y	Y	N	N	Y	Y	Y
10b	T	N	N	Y	N	Y	N	Y	N
11a	V	Y	Y	Y	Y	N	Y	Y	Y
12a	T	Y	Y	Y	N	N	Y	Y	N
12b	V	Y	Y	Y	Y	N	Y	Y	Y
13a	V	Y	Y	Y	N	Y	Y	Y	N
14a	V	N	N	Y	N	Y	Y	Y	N
14b	T	Y	Y	Y	N	N	Y	Y	N
15a	V	Y	N	Y	N	Y	Y	Y	N
16a	V	Y	Y	Y	Y	Y	Y	N	N
17a	V	Y	Y	Y	Y	Y	Y	Y	Y

Case code	2015 int (visit/tele int)	2016 int	2017 int	Coordinator survey 1	Coordinator survey 2	NCETM report Y1	NCETM report Y2	Pupil survey Year 1	Pupil survey year2
18a	V	Y	Y	N	N	Y	Y	N	N
18b	T	Y	N	Y	N	N	N	N	N
19a	T	Y	Y	Y	N	Y	Y	Y	N
19b	V	Y	Y	Y	Y	N	Y	N	N
20a	V	Y	Y	Y	Y	Y	Y	N	Y
21a	V	Y	Y	Y	Y	Y	Y	Y	N
22a	T	Y	Y	Y	Y	Y	Y	Y	Y
22b	V	Y	Y	Y	Y	Y	Y	Y	Y
23a	T	Y	Y	Y	N	Y	N	N	N
24a	T	Y	Y	Y	N	N	Y	N	N
25a	V	Y	Y	Y	Y	N	N	Y	N
25b	T	Y	Y	Y	N	Y	Y	N	N
26a	V	N	Y	Y	N	Y	N	Y	N
27a	V	Y	Y	Y	Y	Y	Y	Y	Y
27b	T	Y	Y	Y	Y	N	Y	Y	Y
28a	T	Y	Y	Y	Y	Y	Y	Y	Y
29a	V	Y	Y	Y	Y	Y	Y	Y	Y
29b	T	Y	Y	Y	N	N	Y	Y	N
28b	V	Y	N	N	N	N	Y	Y	N
30a	V	Y	Y	Y	N	Y	Y	Y	N
31a	V	Y	Y	Y	N	Y	Y	Y	Y
31b	V	Y	Y	N	Y	Y	Y	Y	Y
32a	V	N	N	Y	N	N	N	N	N
32b	T	Y	N	Y	N	N	N	Y	N

Case code	2015 int (visit/tele int)	2016 int	2017 int	Coordinator survey 1	Coordinator survey 2	NCETM report Y1	NCETM report Y2	Pupil survey Year 1	Pupil survey year2
Total = 48	Total Visit=31	total yes =43	total yes =40	total yes= 45	total yes=22	total yes =34	total yes =37	total yes=34	total yes = 18

6. Strand one supporting data

In this section, additional data and findings are reported that supplement or support findings in the main report.

6.1 School and lead teacher interviewee characteristics

Data presented below are based on interviews with MTE cohort 1 teachers in 2016 and 2017. The number of interviews varied and therefore 2016 data are based on 43 interviews and 2017 data are based on 40 interviews. Consequently, the data are not directly comparable, as the schools involved varied slightly year on year, for example two schools were unable to commit to an interview in year 2 but did take part in year 3.

Of the 40 interviewees in 2017, 26 were the same as the person interviewed in 2016, and just over half (n=21) had been on the exchange visit in 2015. Of the 19 who did not go on the visits, 10 were maths leads, five were assistant head teachers, three were classroom teachers and one was a head teacher.

Table 5: MTE cohort 1 Interviewee job roles 2017

Maths lead	Senior leader	Teacher
19	18	3
48%	45%	8%

Source: MTE cohort 1, interviews 2017 *Percentages may add up to more than 100 due to rounding.

Table 5 above shows that the majority of interviewees in the third year were maths leads (n=19), followed by senior leaders (n=18). A number of maths leads and senior leaders were also class teachers.

Table 6 shows the characteristics of the schools involved.

Table 6: MTE cohort 1 school characteristics

MTE cohort 1 school characteristics	
Teaching school	15
Maths Hub lead school	15
Part of a multi-academy trust which includes a Teaching School	18
Part of a multi-academy trust that includes a Maths Hub lead school	17

Source: MTE cohort 1, interviews 2017

MTE cohort 1 interviewees were asked in qualitative interviews about implementation of mastery in all year groups. Table 7 focuses on implementation in Year 4 and Year 6. This gives a picture of Year 6 classes which had experienced two full years of mastery (n=17).

Table 7: Level of implementation of mastery in MTE cohort 1 schools

	No mastery Y5	Partial mastery Y5	Full mastery Y5	unknown Y5
No mastery Y6	0	4	0	1
Partial mastery Y6	1	9	0	1
Full mastery Y6	0	6	17	0
Unknown Y6	0	4	1	0

Source: MTE cohort 1, interviews 2016 and 2017.

In order to corroborate data in strand two of the research, interviewees were asked if all Year 2 and Year 6 pupils in their school had experienced at least two full years of *substantial* teaching for mastery. Table 8 below provides this data, and shows that Year 2 pupils had experienced substantial teaching for mastery for a full two years in 26 of the 40 schools participating in 2017. This proportion was lower for Year 6 classes where 17 schools had substantially implemented the mastery approach for the full two years. Interviewees were asked to elaborate on their answers. The reasons given for Year 6 classes having been particularly less likely to have had substantial implementation were: the focus on teaching pupils for standard assessment tests (SATs) and/or new teachers joining who were less experienced in the mastery methods of teaching. Some teachers said that they were confident that many of their Year 6 classes had experienced 'some' teaching for mastery over the two years, but did not feel this had been 'substantial' for these reasons.

Table 8: Implementation of substantial teaching for mastery in Years 2 and 6 MTE cohort 1 schools

	Yes		No		unknown	
	N	%	N	%	N	%
All Year 2 pupils have had substantial teaching for mastery	26	54%	14	29%	6	13%
All Year 6 pupils have had substantial teaching for mastery	17	35%	22	46%	8	17%

Source: MTE cohort 1, interviews 2016 and 2017. Two schools are a junior school and therefore do not have a Year 2; one school was an infant school and therefore does not have a Year 6.

Percentages given out of 48 schools

Where data are unknown, this is because an interview was not conducted with a representative from the school.

6.2 Changes to practice 2016 and 2017 in MTE cohort 1 schools

Use of representations

Figure 5 provides descriptions of different levels of use of representations and is reproduced from the third interim report. Tables 9 and 10 present analysis of levels of use of representations in the MTE cohort 1 schools sampled for interviews in 2016 and 2017 respectively. These data informed the analysis presented in the main report section 5.

Figure 5: Use of visual and concrete representations

	Visual representations	Concrete representations
Limited	Using more visual aids such as photographs or clip art, but not linked to mathematical models or mathematical learning; or having intentions to introduce greater use in the future; more mathematically meaningful practices only appeared to be happening in the lead primary teachers' classes.	Used with younger learners or low attaining pupils and either did not refer to specific materials of such references are limited. Typically materials such as dienes blocks or counters are used for modelling addition or subtraction and simple arithmetic only. In some cases interviewees referred to intentions, or increased awareness rather than to changed practice.
Embedding	Increasing use of visual representation as mathematical models; aiming for consistency in every lesson; some reference made to challenge for some teachers; the Concrete-pictorial-abstract approach was mentioned by some as something that was being adopted. The bar model was frequently referred to as one specific example.	Increasing use including more use in KS2 and across the attainment range, but use inconsistent; more equipment purchased to give access to all classes or 'getting it out of the back of the cupboard'; references made to the concrete-pictorial- abstract with examples of concrete representation as the start of a topic; reporting that teachers are developing knowledge of how to use these with all years and a wider variety of mathematical content.

	Visual representations	Concrete representations
Embedded use	Multiple and varied visual models used and linked mathematically; different forms of representation were linked, for example referring to concrete-pictorial-abstract as a triangle or to be used alongside each other rather than a sequence; use of models linked to other practices such as questioning, or variation theory; some schools had formulated the approach in policy, for example, to always use two representations in every lesson; patterns of use are consistent across the school.	Used in every lesson and/or across whole school and/or full attainment range; a wide variety of materials are discussed with reference to appropriateness for different mathematical content; routinely, concrete materials are on desk for students to use during explanation; the importance of moving between different representations was discussed, and referring to concrete-pictorial-abstract as a triangle or to be used alongside each other; some discussed creating their own specialised concrete materials for particular topics.

Frequency of use of representations

Table 9: Frequency of types of use of representations 2016 (n=43)

		Visual			
		Limited	Embedding	Embedded	**Total**
Concrete	Limited	2	8	1	**11**
	Embedding	2	14	4	**20**
	Embedded	0	3	9	**12**
	Total	**4**	**25**	**14**	

Source: MTE cohort 1, interviews 2016

Table 10: Frequency of types of use of representation 2017 (n=38)

		Visual			
		Limited	Embedding	Embedded	**Total**
Concrete	Limited	2	1	0	**3**
	Embedding	1	13	4	**18**
	Embedded	0	2	15	**17**
	Total	**3**	**16**	**19**	

Source: MTE cohort 1, interviews 2017

Note that for the two schools which participated in the 2017 interviews but not the 2016 interviews there was insufficient data to categorise their use of representations.

Promoting conceptual and procedural fluency at school level

In 2017, interviewees were asked about their approaches to promoting conceptual understanding and procedural fluency. After outlining the ways in which teachers in their school achieved this, they were asked whether these approaches were an individual teacher's choice or an expectation across the school or school policy. Responses are summarised in Table 11. Although many interviewees answered that there was a mix of formal expectation and teacher discretion (n=13), the majority of teachers said that the approaches outlined were an expectation across the school or school policy (n=24). This indicates that strategies to promote conceptual understanding and procedural fluency were being embedded throughout the school, not simply advised as best practice or used only by those teachers most experienced with teaching for mastery.

Table 11: Approaches to promoting conceptual understanding and procedural fluency 2017

An expectation across the school	School policy	Teacher choice	Mixed
18	6	3	13
45%	15%	8%	33%

Source: MTE cohort 1, interviews 2017

Use of textbooks

Although textbooks were used for a variety of purposes by the majority of MTE cohort 1 schools in 2017 (n=29), the extent and type of use varied between schools as Table 12 indicates. Some teachers for example used textbooks for planning only (n=9). Interviews revealed that the use of textbooks varied between year groups also.

Table 12: Use of textbooks in 2016 and 2017

Use of textbooks	No. of schools 2016	No. of schools 2017
Use with children	8	1
Used for planning only	9	9
Used for planning for some year groups/classes and with children in others	8	19

Source: MTE cohort 1, interviews 2016 and 2017

Lesson activities

As outlined in the interim reports, a typical English mathematics lesson would involve a three-part structure of: teacher input, pupil practice and a final plenary. In contrast, Shanghai mathematics lessons are made up of multiple short activities based on more whole-class teacher-pupil interaction. In an attempt to gauge the extent to which teachers were adopting this Shanghai style of lesson structure, in 2017 interviewees were asked how many different changes in activities happened during the course of a typical mathematics lesson. Most teachers (n=23) answered that there were between four and six changes in activity, compared to only two teachers reporting only one to three changes in activity. This indicates a partial move away from the standard three-part lesson, but does not necessarily suggest a move to a Shanghai-style structure. Fifteen teachers, however, answered that they have more than six changes in activity during a lesson. Table 13 summarises these data.

Table 13: Number of different activities in a typical mathematics lesson

1-3	4-6	7-9	10-12	12 +
2	23	10	4	1
5%	58%	25%	10%	3%

Source: MTE cohort 1, interviews 2017

* Percentages may add up to more than 100 due to rounding.

Differentiation

Teachers were asked a number of questions about differentiation in the second round of interviews. As part of the analysis process for the 2016 interviews, teachers' answers to these questions were categorised as one of three different approaches to differentiation. These categories are specified in Table 14 below. During the 2017 interviews, teachers were asked to confirm whether they felt they had been categorised correctly and if anything had changed in their approach to differentiation. As can be seen in Table 16, the majority of interviewees had been categorised as, or felt that their approach now was in line with, 'Differentiation by deepening and support'. This is a substantial departure from an average English mathematics lesson where teachers would be expected to provide perhaps four different tasks differentiated according to their perceptions of pupils' abilities, with the most difficult tasks moving pupils on to a higher level or even a different topic area. Differentiation by deepening and support provides more challenging tasks for those who have successfully completed the core activity, but through depth activities instead of moving on. Those pupils who find the core task difficult are supported and possibly given intervention to support learning.

Table 14: Approaches to differentiation 2016 and 2017

Differentiation	No. of schools 2016	No. of schools 2017
Differentiation by deepening and support	31	32
Differentiation in transition	5	4
Differentiation by allocated task	7	7

Source: MTE cohort 1, interviews 2016 and 2017

Grouping

In the first round of data collection, survey data showed how MTE cohort 1 schools were grouping their pupils, indicating the percentage of schools grouping by class, within class or not grouping by ability at all (heterogeneous grouping). More information was then collected on grouping via qualitative interviews in 2016 and 2017. The data on grouping from all three years are presented in Table 15 below.

Table 15: Grouping arrangements for pupils in mastery classes 2016 and 2017

Form of grouping	Percentage of schools 2015 (majority of all classes)	Percentage of schools 2016 (substantial mastery classes)	Percentage of schools 2017 (substantial mastery classes)
Heterogeneous grouping (pupils not set or grouped by attainment within class)	38%	67%	70%
Pupils set by class (pupils allocated to classes based on prior attainment and/or perceptions of 'ability')	22%	14%	10%
Pupils grouped by prior attainment within class (pupils of similar attainment sat together)	40%	19%	13%
Mixed across year groups or classes	N/A	N/A	8%

Source: MTE cohort 1 2015 survey; interviews 2016, 2017

As can be seen from Table 15 grouping arrangements in classes where MTE mastery pedagogy is implemented are considerably different from the usual practices in schools prior to the MTE. Although the figures are not strictly comparable year to year due to a slight variation in respondents, the increase in heterogeneous grouping from 38% to 70% over the three-year period represents a substantial change in practice.

There had been a slight increase in schools reportedly moving away from grouping pupils by attainment from 2016 to 2017. A caveat to note here is that for a small number of schools, despite choosing not to group pupils by ability in most year groups, grouping was still happening in Year 6. This was said to be due to SATs preparation or a large gap in ability already in place in this year group from previous setting arrangements. A small number of teachers explained how the classes/years in the school were sometimes split whereby half would be grouped by ability and half would not. Other teachers simply stated that practice would vary around the school dependent on the class and the teacher. (MTE cohort 1, school 16A, interview 2017).

The rationale for schools to move away from grouping was stated by some teachers to be linked to their knowledge of empirical evidence outlining the benefits for pupils of being in heterogeneous groups:

The research says that setting isn't successful; it isn't the right thing to do. (MTE cohort 1, school 18A, interview 2017)

Basically the Sutton Trust and their research, all different academic research shows that streaming doesn't work. (MTE cohort 1, school 22A, interview 2017)

Teachers talked about how grouping did not fit into a mastery approach to teaching and felt that they could not rationalise continuing to group pupils while teaching for mastery:

It just became really apparent that what we were modelling to the children was neither growth mind-set nor mastery. (MTE cohort 1, school 22A, interview 2017)

This was also related to the potential impact on pupils of being put into a lower ability group:

You're saying to a child at the age of six, 'You're not very good at maths. To give a child that impression at the age of six is completely wrong. It's about giving aspiration to all of them so they've all got the same. (MTE cohort 1, school 18A, interview 2017)

Similarly to the previous year's findings, there was a shift in mind-sets about how pupils learn mathematics. Some interviewees articulated this as a move to viewing pupils as learners who may have greater ability in some aspects of learning mathematics than other aspects, and therefore no particular pupils should be labelled as higher or lower ability. Having heterogeneous groups, and sometimes sitting pupils in mixed-ability/attainment pairs or groups, was said to facilitate learning by exposing pupils to all aspects of learning and having maths 'modelled' for them by a learning partner. This enabled certain pupils to access areas of the curriculum which were effectively barred from them previously because they were designated to a lower ability group. Some teachers described how they felt the approach was narrowing the gaps, but was also providing appropriate challenge to their higher-attaining pupils due to them being asked to think deeply about concepts and articulate their understanding, as the example below describes:

We've done some work on slow thinking and fast thinking. Being fast isn't what makes good maths. That's the fluency issue. Our children who are fluent think that makes them good at maths, because they think that's all that good maths is about. (MTE cohort 1, school 31b, interview 2017)

Seating

Similarly to the reports in 2016, the majority of MTE leads in 2017 (n=19) stated that for most mathematics lessons across their school, pupils were seated in small groups (see Table 16 below). In addition to questions about seating arrangements, in 2017 interviewees were asked who decides how children are seated in mathematics classes. Table 17 below conveys that for the majority of schools the class teacher decides seating arrangements.

Table 16: Seating arrangements 2016 and 2017

Seating	No. of schools 2016	No. of schools 2017
Small groups	22	19
Rows	12	8
Other/mixed	9	13

Source: MTE cohort 1, interviews 2016 and 2017

Table 17: Who decides how children are seated in mathematic classes in 2017

Teacher	School policy	Year group leader	Pupil
27	10	1	2
68%	25%	3%	5%

Source: MTE cohort 1, interviews 2017

Intervention

Interviewees were asked to confirm whether or not their intervention practices were the same as the previous year. The three tables below provide details for 2016 and 2017. In terms of daily intervention, practices appeared to have remained the same or very similar, as 29 interviewees (as for the last year) stated that intervention happens on a daily basis. The roles of staff who deliver the interventions appears to be fairly consistent with the previous year, with a mix of teachers and TAs being the most likely answer (n=20). Similarly, information about how frequently pupils needing additional support are identified was broadly similar to the last year, with identification happening daily in the vast majority of schools (n=33).

Table 18: Frequency of intervention and frequency of pupil identification for intervention 2016 and 2017

Frequency of intervention	No. of schools 2016	No. of schools 2017
Daily	29	29
Less than daily	15	9
Unknown	0	2

Frequency of pupil identification	No. of schools 2016	No. of schools 2017
Daily	31	33
Less than daily	12	6
Unknown	0	1

Source: MTE cohort 1, interviews 2016 and 2017

Table 19: Staff working with pupils during intervention 2016 and 2017

Staff	No. of schools 2016	No. of schools 2017
TA	6	6
Teacher	16	17
Teacher and/or TA	20	20

Source: MTE cohort 1, interviews 2016 and 2017

2017 saw a slight increase in the number of schools which had split mathematics lessons (13) up from 11 the previous year. The majority of schools (n=28) still have one mathematics lesson a day and only a small proportion (n=7) have two lessons. Consistent with last year, the time spent on mathematics each day was just over an hour (64 minutes), with 22 of the 40 schools having 60-minute lessons.

Table 20: Split lessons 2017

Lessons are split into two parts	Lessons are not split
13	27
32%	68%

Source: MTE cohort 1, interviews 2017

Timetabling

Table 21 presents data on timetabling of mathematics lessons and interventions. It is notable that there is a slight decrease in the number of schools undertaking intervention after a lesson. In interviews, a number of participants indicated that their schools had experimented with same-day intervention after the lesson but had reverted back to previous practices. Nevertheless, generally the approach to intervention in schools had changed as a result of engagement in the MTE programme.

Table 21: Timing of intervention 2016 and 2017

Timing of intervention	No. of schools 2016	Percentage 2016	No. of schools 2017	Percentage 2017
After lessons	26	60.5	21	52.5
Before lessons	1	2.3	0	0.0
During lessons	9	20.9	13	32.5
Mixed	7	16.3	6	15.0

Source: MTE cohort 1, interviews 2016 and 2017

Table 22 presents data on the number of mathematics lessons per day in MTE cohort 1 schools.

Table 22: Number of mathematics lessons in a day 2017

1 lesson per day	1 lesson + mental arithmetic	2 lessons per day	2 lessons + mental arithmetic
28	5	4	3
70%	13%	10%	8%

Source: MTE cohort 1, interviews 2017 *Percentages may add up to more than 100 due to rounding.

Table 23: Time spent on mathematics on an average day in 2017

	Time
Mean	66 minutes
Median	60 minutes
Mode	60 minutes
Range	45 minutes

Lesson preparation and design

Interviewees were asked in more depth in 2017 about how lessons are prepared than in the 2016 interviews. They were asked how often they plan mathematics lessons in pairs or in groups. Consistent with lesson preparation in Shanghai, as Table 24 illustrates, 27 (68%) interviewees reported preparing lessons either in a pair or in a group once a week or more often.

Table 24: Lesson planning in pairs or groups 2017

More than once a week	Once a week	once or twice per half term	Never	Unknown
11	16	7	5	1
28%	40%	18%	13%	3%

Source: MTE cohort 1, interviews 2017 *Percentages may add up to more than 100 due to rounding.

Whether teachers planned in pairs or groups was often said to be related to the school size. Where schools were one-form entry, group or paired planning was seen as more difficult. In schools where teachers did plan with one or more colleagues, they often used their PPA (planning, preparation and assessment) time for this. Some teachers described how planning had become more standardised across the school, and a smaller number talked about working with other schools to prepare lessons, or standardising lesson plans across a trust. Teachers explained that detailed planning might happen for the first couple of lessons in the week and then planning would be adapted or tailored to needs as the week went on, depending on pupils' responses to these first lessons:

Within our PPAs we have a brainstorming map which is where we map out the learning. Then we have daily lesson plans that include things like key questions, vocabulary, resources and support that's given to individual children and to extend, but those plans are done by individual teachers. Every class will access it at a different rate of learning. Also those are annotated on and adapted as the week goes on. (MTE cohort 1, school 23A, interview 2017)

Interviewees were asked about whether planning had changed. Responses were mixed, with some teachers said that the way they planned was the same, but the content of the lessons being planned for had changed greatly. However, other teachers felt that their planning had changed dramatically as a result of the exchange. Teachers talked mostly about how they were now planning for understanding instead of simply planning the activities to be worked through. This was explained as a change from a procedural approach to planning, focusing on 'how we teach', to a conceptual approach centred on how pupils learn:

Move away from what the children are going to do, to thinking about what the children are going to learn.' (MTE cohort 1, school 31b, interview 2017)

This approach was seen by some teachers to be facilitated by their school having moved away from setting or traditional approaches to differentiation.

Teachers cited the White Rose and Maths No Problem schemes as the main resources they would go to when planning. However a number of teachers described how they would pick out different aspects from a variety of resources to create their own materials. What constituted a 'lesson plan' varied: some teachers were planning in greater detail than they did previously, while other teachers saw the lesson plans more as a guide with notes to support teaching than as something they would stick to rigidly:

I literally have moved away from a planning proforma, because I felt that moved teachers towards what are the children going to do? So we just go to a big blank piece of A3 paper and they draw an S on it and they just plan the journey from that. (MTE cohort 1, school 31b, interview 2017)

6.3 Implementation pathways

In the main report (section 7.8) a model of implementation pathways was included. Table 25 below provides detail of the analysis that supports that discussion. Note that in order to preserve anonymity school codes are not included. The categorisations of 'full' and 'mastery' refer to implementation of MTE mastery pedagogy.

Table 25: Implementation pathways table

Pathways			Frequency
Start	**Pathway**	**Categorisation**	
Already committed and mastery textbook/scheme	Mastery textbook/scheme	Full	1
Already committed and mastery textbook/scheme	Mastery textbook/scheme	Mastery	2
Already committed and mastery textbook/scheme	Mastery textbook/scheme and PMTMSP	Full	1
Already committed and mastery textbook/scheme	Mastery textbook/scheme and PMTMSP	Mastery	1
Already committed		Full	1
Already committed	mastery textbook/scheme	Mastery	1
Already committed	mastery textbook/scheme and PMTMSP	Full	1
Already	PMTMSP	Not mastery	1
Newly committed		Full	6
Newly committed	PMTMSP	Full	2
Newly committed	Mastery textbook/scheme	Full	6
Newly committed	Mastery textbook/scheme and PMTMSP	Full	6
Newly committed	Textbook	Mastery	3
Newly committed	Mastery textbook/scheme and PMTMSP	Mastery	1
Newly committed	PMTMSP	Not mastery	1
Newly committed		No full data	7
Newly committed		Not mastery	1
Newly committed	Textbook	No full data	3
Cautious	Textbook	Low	1
Cautious	Textbook	Mastery	1
Cautious	mastery textbook/scheme and PMTMSP	Full	1

Note that the attribution of commitment was based on checking judgements made on the basis of the 2015 interviews, during the 2016 interviews. If a 2016 interview was not completed then the 2015 researcher judgement was used.

6.4 Pupil outcomes

Table 26 (below) shows that the majority of MTE leads felt that pupils had progressed more than expected.

Teachers were more confident that their pupils' attainment had been influenced positively by the mastery approach to teaching. In 2016, 18 teachers answered that they perceived their pupils' attainment to be more than expected; in 2017 this had increased to 27 teachers (despite the lesser number of interviewees). This suggests that the impact of the exchange on pupils has increased as schools have had a longer period of time in which to embed the approach across the school.

Table 26: Perceptions of pupil outcomes 2016 and 2017

Perception of pupil outcome statements	More than expected	About as expected	Below expected	Not able to answer
Pupils' attainment. N 2016	18	14	1	10
Pupils' attainment. N 2017	27	3	0	10

Source: MTE cohort 1, interviews 2016 and 2017

6.5 Mathematics Coordinator Survey 2017: supplementary data

Survey responses from the 2017 Mathematics Coordinator Survey were firstly looked at stratified by Cohort (1&2) and by PMTMS involvement (1&2). However, stratification by these components yields unequal and small sample sizes. Response data was therefore re grouped. Group 1, those respondents reporting being in cohort one or two and/or having undertaken the PMTMS training, group 2, those respondents not involved in an MTE cohort or PMTMS but reporting substantial mastery and group 3, those not involved in an MTE cohort or PMTMS training and NOT reporting substantial mastery. The below tables show the percentage of responses to each question answer stratified by group. Kruskal Wallis (for ordinal variables) and Chi-Square tests (for nominal variables) for significance were ran to determine whether differences between groups were significant, but do not illustrate where the difference lies. Statistical test results are included in parentheses after each survey question, a p-value of <0.05 is considered significant.

Table 27: Full curriculum access for all mastery component

Full curriculum access for all			
	Group 1 (MTE and/or PMTMS cohort	**Group 2 (No MTE and/or PMTMS but reported substantial mastery)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery)**
Differentiated tasks are set for pupils (X2 (2)=81.82, P<0.01)			
Always	4.0	10.2	17.6
Often	16.2	27.6	41.2
Sometimes	22.0	22.4	25.5
Rarely	43.4	30.6	13.7
Never	14.5	9.2	2.0

Full curriculum access for all			
	Group 1 (MTE and/or PMTMS cohort	Group 2 (No MTE and/or PMTMS but reported substantial mastery)	Group 3 (No MTE and/or PMTMS and no substantial mastery)
Main activities would be the same, differentiation by outcome (X2 (2)= 20.17, P<0.01)			
Always	17.4	13.7	6.8
Often	46.1	36.8	35.4
Sometimes	27.5	34.7	40.8
Rarely	8.4	12.6	16.5
Never	0.6	2.1	0.5
Pupils learn the main content first, tasks are set to deepen understanding (X2 (2)= 35.75, P<0.01)			
Always	55.9	59.8	32.2
Often	35.5	35.1	47.6
Sometimes	5.9	5.2	17.3
Rarely	1.6	0.0	1.9
Never	1.1	0.0	1.0
Frequency of identification of additional support (X2 (2) 18.49, P<0.01)			
Daily	45.7	41.8	28.2
A number of times per week	38.3	31.6	39.9
Weekly	5.3	10.2	12.2
Half termly	9.0	15.3	14.1
Less often than half termly	1.6	1.0	5.6

Full curriculum access for all			
	Group 1 (MTE and/or PMTMS cohort	Group 2 (No MTE and/or PMTMS but reported substantial mastery)	Group 3 (No MTE and/or PMTMS and no substantial mastery)
Who gives additional support (X2 (4)=11.95, P<0.05)			
Teacher	11.7	5.1	6.1
TA	11.2	20.4	21.2
Teacher and a TA	77.1	74.5	72.6
Amount of curriculum covered in the last three years (X2(2)=55.65, P<0.01)			
Increased	27.8	37.4	59.1
Stayed the same	26.2	27.3	28.0
Decreased	46.0	35.4	13.0

Table 28: Varied interactive teaching mastery component

Varied Interactive Teaching			
	Group 1 (MTE and/or PMTMS cohort	Group 2 (No MTE and/or PMTMS but reported substantial mastery)	Group 3 (No MTE and/or PMTMS and no substantial mastery)
Whole class teacher-pupil interaction (X2(2)=41.66, P<0.01)			
Increased	73.3	53.1	39.8
Decreased	20.9	32.7	45.4
Stayed the same	5.9	14.3	14.8

Varied Interactive Teaching			
	Group 1 (MTE and/or PMTMS cohort	**Group 2 (No MTE and/or PMTMS but reported substantial mastery)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery)**
Lesson structure (X2(2)=63.19, P<0.01)			
Starter, introduction, activity, teacher explanation, practice then plenary	22.0	38.4	62.2
Multiple periods of questioning, teacher pupil dialogue, pupils working on 1/2 problems/tasks	78.0	61.6	37.8
Teacher-pupil interaction frequency (X2(2)=25.51, P<0.01)			
Increased	82.9	79.8	61.1
Decreased	15.5	18.2	34.7
Stayed the same	1.6	2.0	4.1

Table 29: Knowledge of mathematical facts and Language mastery component

Knowledge of Mathematical Facts and Language			
	Group 1 (MTE and/or PMTMS cohort	**Group 2 (No MTE and/or PMTMS but reported substantial mastery)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery)**
Precise mathematical language by teachers and pupils (X2(2)=27.53, P<0.01)			

Knowledge of Mathematical Facts and Language			
	Group 1 (MTE and/or PMTMS cohort	**Group 2 (No MTE and/or PMTMS but reported substantial mastery)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery)**
Always	31.6	19.2	12.5
Often	51.9	63.6	54.9
Sometimes	16.6	16.2	27.2
Rarely	0.0	1.0	5.4
Never	0.0	0.0	0.0
Key ideas and concepts recited individually or as a class (X2(2)=28.52, P<0.01)			
Always	13.4	8.1	5.4
Often	44.1	41.4	28.6
Sometimes	35.5	40.4	44.3
Rarely	6.5	8.1	18.4
Never	0.5	1.0	2.2
Not sure	0.0	1.0	1.1
Teachers ask for explanations about how answers were obtained (X2(2)=30.51, P<0.01)			
Always	55.1	42.4	27.9
Often	40.6	54.5	61.2
Sometimes	4.3	3.0	9.8
Rarely	0.0	0.0	0.5
Never	0.0	0.0	0.0
Not sure	0.0	0.0	0.5

Knowledge of Mathematical Facts and Language			
	Group 1 (MTE and/or PMTMS cohort	Group 2 (No MTE and/or PMTMS but reported substantial mastery)	Group 3 (No MTE and/or PMTMS and no substantial mastery)
Pupils are encouraged to communicate mathematically to the whole class (X2(2)=19.76, P<0.01)			
Always	32.6	16.2	18.0
Often	54.5	64.6	53.6
Sometimes	10.7	19.2	26.2
Rarely	2.1	0.0	2.2
Never	0.0	0.0	0.0

Table 30: Mathematical and meaningful coherent activity mastery component

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)	Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)
Multiple representations for a mathematical concept or procedure are used in a single lesson (X2(2)=63.38, P<0.01)			
Always	28.5	22.2	5.1
Often	55.4	47.5	45.8
Sometimes	16.1	27.3	41.8
Rarely	0.0	3.0	5.6
Never	0.0	0.0	1.7

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Concrete materials are used with all year groups in the school (X2(2)=27.90, P<0.01)			
Always	27.8	33.3	16.4
Often	55.6	47.5	44.6
Sometimes	14.4	16.2	28.8
Rarely	2.1	3.0	10.2
Never	0.0	0.0	0.0
Concrete materials are used with pupils of all attainment levels in the school (X2(2)=44.54, P<0.01)			
Always	30.5	35.4	14.1
Often	50.8	46.5	39.0
Sometimes	15.5	14.1	35.0
Rarely	3.2	3.0	11.3
Never	0.0	0.0	0.0
Not sure	0.0	1.0	0.6

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Concrete/pictorial and symbolic representations are used together in all lessons (X2(2)=58.09, P<0.01)			
Always	26.9	29.3	9.6
Often	52.7	52.5	38.4
Sometimes	18.8	15.2	39.5
Rarely	1.6	2.0	12.4
Never	0.0	0.0	0.0
Not sure	0.0	1.0	0.0
The sequence of concrete-pictorial abstract forms of representation is used to structure teaching (X2(2)=48.46, P<0.01)			
Always	23.7	32.3	10.2
Often	54.8	44.4	37.9
Sometimes	17.7	21.2	39.5
Rarely	2.7	1.0	10.7
Never	1.1	0.0	0.6
Not sure	0.0	1.0	1.1

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Teachers move back and forth between different forms of representation in order to connect them and support understanding (X2(2)=63.66, P<0.01)			
Always	25.5	26.3	5.6
Often	52.7	47.5	39.0
Sometimes	20.2	23.2	42.4
Rarely	1.1	2.0	11.9
Never	0.0	0.0	0.0
	0.5	1.0	1.1
Using representations and models to introduce concepts (X2(2)=44.41, P<0.01)			
Always	41.2	42.4	16.4
Often	50.8	50.5	58.5
Sometimes	7.5	7.1	24.0
Rarely	0.5	0.0	1.2
Never	0.0	0.0	0.0

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Using teacher questioning and classroom dialogue (X2(2)=21.51, P<0.01)			
Always	57.8	64.6	40.4
Often	40.6	35.4	52.6
Sometimes	1.6	0.0	6.4
Rarely	0.0	0.0	0.6
Never	0.0	0.0	0.0
Making understanding an explicit focus in lesson preparation (X2(2)=32.14, P<0.01)			
Always	48.1	47.5	23.8
Often	44.4	47.5	58.1
Sometimes	7.5	5.1	16.3
Rarely	0.0	0.0	1.7
Never	0.0	0.0	0.0

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Starting from a problem or carefully selected task (X2(2)=30.06, P<0.01)			
Always	27.3	25.3	4.4
Often	39.0	37.4	42.1
Sometimes	28.9	36.4	48.0
Rarely	4.8	1.0	5.3
Never	0.0	0.0	0.6
Connecting different mathematical concepts and procedures (X2(2)=35.56, P<0.01)			
Always	29.4	23.2	9.9
Often	49.7	59.6	48.0
Sometimes	19.3	16.2	38.0
Rarely	1.1	1.0	4.1
Never	0.0	0.0	0.0
Not sure	0.5	0.0	0.0

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Emphasising the precise use of Mathematical language (X2(2)=29.31, P<0.01)			
Always	42.2	45.5	24.6
Often	47.1	47.5	48.0
Sometimes	10.2	7.1	24.0
Rarely	0.5	0.0	3.5
Never	0.0	0.0	0.0
Planning for and addressing misconceptions (X2(2)=41.51, P<0.01)			
Always	42.2	48.5	17.1
Often	47.1	42.4	59.4
Sometimes	10.7	9.1	19.4
Rarely	0.0	0.0	3.5
Never	0.0	0.0	0.6

Mathematically Meaningful and Coherent Activity			
	Group 1 (MTE and/or PMTMS cohort (%)	**Group 2 (No MTE and/or PMTMS but reported substantial mastery) (%)**	**Group 3 (No MTE and/or PMTMS and no substantial mastery) (%)**
Using a textbook, scheme or designed resource (X2(2)=9.92, P<0.01)			
Always	15.6	17.3	5.3
Often	31.7	28.6	32.0
Sometimes	34.9	32.7	34.3
Rarely	15.6	16.3	20.7
Never	2.2	5.1	7.1
Not sure	0.0	0.0	0.6
Using in-class formative assessment (X2(2)=24.93, P<0.01)			
Always	50.3	49.0	26.0
Often	41.2	40.8	57.4
Sometimes	8.0	10.2	15.4
Rarely	0.5	0.0	1.2
Never	0.0	0.0	0.0

7. Strand one: Implementation criteria and analysis

7.1 Process for determining implementation criteria

All MTE cohort 1 schools which contributed interviewees in 2017 described substantial implementation of mastery approaches in at least some classes. However, the detailed responses about practices suggested that implementation levels varied. Consequently, a set of implementation criteria was devised based on data analysis and informed by benchmarking against schools that appeared to have high levels of implementation and which in many cases had participated in the PMTMSP.

Table 31 below specifies the implementation criteria in terms of the four main components of MTE mastery pedagogy and their indicators. Table 32 sets out how sub-components were combined to form the categories of *high mastery, mastery and low/not implemented.* Lastly, Table 33 shows how the four main components were combined to determine the final categorisation for a school. Greater weight was given to component 1 (varied interactive teaching) and component 2 (meaningful and coherent activity) than to the other two components. This reflected the focus in MTE cohort 1 schools (and NCETM's formulation of TfM) on pedagogical and classroom practice with other practices such as intervention policy designed to support this.

To ensure reliability of analysis, an independent researcher without knowledge of the schools was contracted to rate 12 schools separately. They did this by taking a more holistic view, rather than using Nvivo for detailed coding, using the implementation criteria grid as a matrix to shade patterns of implementation based on 2016 and 2017 data.

The outcome of this independent rating was agreement in nine cases. In the other three, the independent researcher had rated one as *high mastery* which the evaluation team had rated as *mastery*. In the other two cases, the independent researcher reported that although there was the appearance of mastery implementation, this was not fully convincing - for example, a school was using a textbook scheme as the main resource that did not align well with a mastery approach. In one of the cases, the school reported that mastery practices were being used only with a small number of classes. Thus, these were rated independently as borderline. The evaluation team had rated these two schools as *low/not implemented*.

This inter-rater checking process indicates that, depending on definitions, the categorisations made by the evaluation team may have under-estimated, in up to four cases, the number of schools in MTE cohort 1 which have implemented mastery approaches to a substantial extent - that is high mastery. However, this does not affect the strand two analysis, as this is based on comparison of the MTE mastery/high mastery schools with contrast schools rather than with schools in the MTE cohort 1 sample categorised in other ways If a less stringent application of criteria was taken and one or more of these found schools were included in the high mastery sample for the impact analysis then this would mean that the size of impact reported in section 8 of the main report would be potentially lower. This is because the impact identified was a little higher for the high mastery sample than for the mastery sample.

Table 31: MTE Mastery pedagogy implementation criteria

	Full MTE mastery pedagogy indicators	MTE Mastery pedagogy implementation indicators	Low or no implementation of component
1a Substantial whole-class varied teaching in multiple part lessons	Multiple part lessons, whole-class teacher-led episodes are a substantial part of the lesson but with interaction, clear sense of different purposes of lessons	Mixture of lesson forms with some multiple part lesson teaching with whole-class as central Multiple part lessons or other if clear that whole-class teaching central	Three-part lessons or unclear about lessons
1b Interactive dialogue	In whole-class teaching episodes use of to-and fro, interaction for substantial portions of time, pupils routinely talk to other pupils, pupils come to the front	To and fro, and similar patterns occur but not necessarily as an essential lesson feature and focus of lesson design, teacher-pupil interaction likely to be consistently happens in lessons; pupils talking to each other reported as a regular feature of lessons.	Lower levels of dialogue. For example responding sometimes or rarely when asked about dialogue practices
2a Depth, meaning, problem solving	Multiple approaches to developing conceptual understanding - using approaches to developing procedural fluency that support conceptual understanding as well e.g. models and representations, reference to conceptual and procedural variation (in a way that shows understanding of this), consistently starting from a problem, using deepening tasks that focus on mathematical structure.	At least one approach to developing conceptual understanding is indicated alongside a conceptual approach to procedural fluency (e.g. using models, or choice of questions - overall three or more), may refer to problem solving but not central to practice	No particular focus on conceptual understanding and procedural fluency or unclear - less than 3 mentioned for developing conceptual understanding and procedural fluency

	Full MTE mastery pedagogy indicators	**MTE Mastery pedagogy implementation indicators**	**Low or no implementation of component**
2b Models and representations	Multiple use of models and representations, not just for low attaining or young children but consistently - might use CPA as a heuristic	Embedding use of models and representations and/or concrete materials - being used more widely, but not yet consistent	Limited implementation in one or both or unclear
2c Mathematically coherent lesson design	Use of East Asian informed materials or equivalent consistently or lesson design consistently that accords with MTE mastery pedagogy - interaction, aim for conceptual understanding and procedural fluency, deepening etc.	Some relationship of materials to mastery, mastery is informing choice Lesson activities informed by MTE mastery pedagogy Use of compatible materials CPA as heuristic, White Rose, NCETM mastery assessment materials, or choice of problems and/or some use of East Asian informed materials for lesson design but not consistent	Other materials or unclear
3a Curriculum pace for whole class access	Slowing curriculum pace	Slowing curriculum pace	No change to curriculum pace
3b Teaching to attainment by deepening and support	All-attainment teaching and differentiation by deepening and support	Setting plus differentiation by deepening and support or not setting with differentiation in advance	Differentiation in advance / give different activities
3c Responsive intervention	Daily intervention and decided daily, intervention outside lessons	Two of three of these aspects present - less frequently than daily, decided weekly, and/or intervention during lessons	Intervention less frequently

	Full MTE mastery pedagogy indicators	MTE Mastery pedagogy implementation indicators	Low or no implementation of component
4a Memorising facts, relationships and structures	Specific times for developing factual knowledge - strategies for this	Specific times for developing factual knowledge - strategies for this sometimes	Not emphasised or discussed
4b Precise mathematical language	Use of precise language - always in lessons, stem sentences	Precise language - sometimes	Not discussed or mentioned

Table 32: Combining sub-components of mastery

Component	High mastery	Mastery implementation	Low/not implemented with indicators
Component 1 combined	1a mastery and 1b full	Mastery in 1a and 1b	Not meeting criteria Does not meet criteria for mastery in 1a and/or 1b or both
Component 2	High - Meets two out of three subcomponents at full	Meets all three at mastery	One or more not met at mastery
Component 3 combined	High - Meets two out of three subcomponents at full criteria, other at medium or full	Mastery - meet all three at medium, or one high, one medium, one low	Not meeting criteria Two low, or one low and two medium
Component 4	High mastery - both present	Mastery - one or other (this due to evidence issue with amount of data)	Not meeting either

Table 33: Determining an overall mastery categorisation

Category	Criteria
High	Components 1 and 2 - both judged as high mastery, and at least one of 3 or 4 high mastery and the other at least mastery
Mastery	Components 1 and 2 - mastery Components 3 and 4 - mastery
No/low implementation	Not meeting the above criteria

7.2 Analysing the relationship between variation in implementation and school characteristics

We have data on 37 of the 47 MTE schools (35 primary and two junior). Of these, a majority are identified as having 'high mastery' (25 schools, 68%) and the vast majority as ‘mastery’ or higher (33 schools, 89%). Only four schools (11%) are identified as having ‘low mastery’.

For the Key Stage 1 (KS1) analyses, schools that are judged to have implemented mastery/high mastery in general and also self-report substantial implementation two years for the Year 2 2017 cohort, will be included in the analyses. Implementation data from 35 of the 44 primary MTE schools[11] were obtained and a majority of these are identified as implementing mastery in both 2016 and 2017 (26 schools, 74%). Among these 26 schools, 19 (73%) are identified as having high mastery, and 24 (92%) with mastery level or higher.

For the Key Stage 2 (KS2) analyses, schools are judged to have implemented a mastery/high mastery level in general and also self-report substantial implementation specifically for two years for the Year 6 2017 cohort will be included in the analyses. Implementation data from 37 of the 47 MTE schools were obtained. A minority of these are identified as implementing mastery in both 2016 and 2017 (16 schools, 43%). Among these 16 schools, 15 (94%) are identified as having high mastery and all 16 with mastery level or higher.

Table 34 uses two further measures of implementation to compare the two mastery thresholds: the percentage of classes identified as having full or partial implementation; and PMTMSP attendance.

Table 34: Levels of mastery & MTE implementation

	Full % mean (sd)	Partial % mean (sd)	PMTMSP n (%)
Complete Sample (n=37 schools)[1]			
High Mastery (n=25)	80.3 (24.05)	17.8 (21.92)	11 (44%)
Mastery or Higher (n=33)	74.8 (29.49)	23.2 (27.38)	14 (42%)
KS1 Analyses - 26 schools implementing in Y2 in both 2016 & 2017			
High Mastery (n=19)	81.9 (25.85)	17.0 (24.20)	8 (42%)
Mastery or Higher (n=24)	78.0 (29.00)	20.5 (27.98)	10 (42%)
KS2 Analyses - 16 schools implementing in Y2 in both 2016 & 2017			
High Mastery (n=15)	85.1 (24.73)	14.9 (24.73)	5 (33%)
Mastery or Higher (n=16)	86.1 (24.14)	13.9 (24.14)	6 (38%)

[11] Excluding the three junior schools (two of which have mastery data)

Note 1: whilst we have implementation data for 37 of the 47 schools, data were only available for 36 schools concerning the proportion of classes exhibiting full/partial implementation in 2016 and 2017.

The measures seem to cross-validate reasonably well, although the difference between high mastery and mastery is fairly small.

The mastery implementation detail was then attached to the main school-level data file for statistical analysis. Using the 2014 KS2 school census data, Table 35 shows a school-level comparison of KS2 attainment, KS1 to KS2 maths value-added scores and %FSM across the four mastery levels.

Table 35: Comparison of samples with differing levels of mastery 2013/14 KS2 school census data

	KS2 Points	KS1-KS2 Maths VA	%FSM
High Mastery (n=25)	29.9 (1.74)	100.7 (1.17)	20.6 (18.47)
Mastery (n=8)	29.1 (1.78)	100.5 (1.52)	31.4 (24.08)
Low Mastery (n=4)	29.9 (0.41)	100.1 (1.10)	11.7 (7.45)
Missing Data (n=10)	29.5 (1.63)	100.4 (1.50)	24.7 (9.08)
Mastery or higher (n=33)	29.7 (1.76)	100.6 (1.24)	23.2 (20.12)

Table 36 presents the pupil-level comparison at KS1. At KS1, in 2014 the mean KS1 maths score is shown alongside two threshold measures: numbers and percentages of pupils exceeding the expected KS1 maths level; and numbers and percentages of pupils meeting/surpassing the expected KS1 maths level. In 2017, following assessment changes in 2016, only the two threshold measures are available.

Table 37 presents the pupil-level comparison at KS2. At KS2, in 2014 the mean KS2 fine points maths score, mean raw KS2 maths score and mean scores in each of the three KS2 maths papers are shown. In 2017, following assessment changes in 2016, the mean KS2 fine points maths score is replaced by a new mean maths points score. The mean raw KS2 maths score along with mean scores in each of the three 2017 KS2 maths papers are also shown.

Table 36: KS1 Comparison of samples with differing levels of mastery NPD Pupil Level Data, 2014 & 2017

2014	KS1 Maths Score	Exceed	Expect
High Mastery	16.8 (3.50)	370/1,190 (31.1%)	1,132/1,190 (95.1%)
Mastery	16.4 (3.59)	101/396 (25.5%)	369/396 (93.2%)
Low Mastery	17.3 (3.29)	63/180 (35.0%)	175/180 (97.2%)
Missing Data	16.1 (3.61)	101/440 (23.0%)	393/440 (89.3%)
Mastery or higher	16.7 (3.52)	471/1,586 (29.7%)	1,501/1,586 (94.6%)
2017	**KS1 Maths**	**Exceed**	**Expect**
High Mastery	n/a	298/935 (31.9%)	777/935 (83.1%)
Mastery	n/a	75/394 (19.0%)	307/394 (77.9%)
Low Mastery	n/a	42/160 (26.3%)	141/160 (88.1%)
Missing Data	n/a	72/412 (17.5%)	295/412 (71.6%)
Mastery or higher	n/a	373/1,329 (28.1%)	1,084/1,329 (81.6%)

Table 37: KS2 Comparison of samples with differing levels of mastery NPD Pupil Level Data, 2014 & 2017

2014	Fine Points Score	Raw KS2 Maths Score
High Mastery	5.08 (0.858)	76.0 (18.08)
Mastery	4.93 (0.871)	72.4 (19.74)
Low Mastery	5.06 (0.645)	76.3 (14.00)
Missing Data	5.00 (0.904)	74.6 (17.77)
Mastery or higher	5.04 (0.863)	75.1 (18.56)

2017	Maths Points Score	Raw KS2 Maths Score
High Mastery	106.3 (7.59)	81.6 (22.26)
Mastery	104.6 (8.53)	76.0 (25.76)
Low Mastery	104.7 (6.68)	76.5 (21.16)
Missing Data	104.4 (7.56)	74.9 (24.09)
Mastery or higher	105.8 (7.89)	80.0 (23.41)

2014	Arithmetic	Paper 2	Paper 3
High Mastery	15.8 (4.13)	30.5 (7.89)	29.6 (7.30)
Mastery	15.5 (4.08)	29.3 (8.61)	27.6 (8.12)
Low Mastery	16.0 (3.25)	30.7 (6.02)	29.5 (6.19)
Missing Data	15.7 (3.70)	29.9 (7.85)	29.1 (7.40)
Mastery or higher	15.7 (4.12)	30.2 (8.09)	29.1 (7.56)

2017	Arithmetic	Paper 2	Paper 3
High Mastery	32.7 (7.61)	25.9 (7.82)	23.0 (8.37)
Mastery	32.3 (8.18)	22.9 (9.59)	20.8 (9.62)
Low Mastery	31.8 (6.50)	24.0 (7.67)	20.7 (8.31)
Missing Data	30.6 (8.26)	23.6 (8.61)	20.6 (8.87)
Mastery or higher	32.6 (7.77)	25.1 (8.45)	22.4 (8.78)

8. Strand two analysis

8.1 Effect sizes

For both the descriptive and multilevel analyses discussed below, the difference between the MTE exchange and contrast control school samples is converted into an effect size measure.

When an outcome variable is a continuous scale (i.e. KS1 maths between 2013 and 2015; KS2 maths 2013 to 2017), the mean difference is converted into a Cohen's d effect size statistic.

Cohen's d is a widely used standardised statistic that enables effect sizes to be compared across outcomes on differing scales and across different studies, time points etc. According to the teaching and learning toolkit developed by the Educational Endowment Foundation (EEF[12]), a 'very high impact' is indicated by an effect size of (d=) +0.70 standard deviations or greater; 'high impact' by an effect size between +0.45 to less than +0.70 sds; 'moderate impact' by an effect size between +0.19 to less than +0.45 sds; 'low impact' by an effect size between +0.02 to less than +0.19 sds and below +0.02 sds 'very low or no impact'.

When an outcome variable is a categorical attainment threshold (i.e. whether a pupil reaches or exceeds the KS1 maths expected level 2013 to 2017) the percentage difference between the exchange school and matched samples is converted into a (odds-ratio) effect size statistic.

Odds-ratios are widely used statistics that measure the difference between one percentage and another as a ratio of the odds for these percentages. Conveniently, it is possible to convert odds ratios into Cohen's d effect size statistics (Sanchez-Meca et al 2003[13]). Table 38 compares odds-ratio and Cohen's d effect size statistics within reference to the EEF teaching and learning toolkit.

[13] Sanchez-Meca, J., Marin-Martinez, F. & Chacon-Moscoso, S. (2003) Effect-size indices for Dichotomized Outcomes in Meta-Analysis. Psychological Methods 8(4) pp448-467.

Table 38: Comparing Cohen's d & Odds-Ratio effect size statistics

EEF 'size of effect'		Cohens d	Odds Ratio
+VE	Very High	+0.70 or higher	3.56 or higher
	High	+0.45 to <+0.70	2.26 to < 3.56
	Moderate	+0.19 to <+0.45	1.41 to < 2.26
	Low	+0.02 to <+0.19	1.03 to < 1.41
Zero	Very Low / zero	-0.02 to < +0.02	0.96 to < 1.03
- VE	Low	-0.02< to -0.19	0.96< to 0.71
	Moderate	-0.19< to -0.45	0.71< to 0.44
	High	-0.45< to -0.70	0.44< to 0.28
	Very High	-0.70 or lower	0.28 or lower

8.2 Detail of the impact analysis

Overview

In December 2017, NPD pupil level data was obtained for KS1 & KS2 pupils in exchange or matched contrast control schools for 2013, 2016 and 2017. A previous request obtained pupil level KS2 data for 2014 and 2015. In total, the pupil level KS1 and KS2 impact analyses covers five academic years between 2013 and 2017. The impact analyses compares the maths attainment of pupils in exchange schools with the attainment of pupils in the matched schools over this period; 2015 to 2017 representing the three years following the start of the exchange and 2013 & 2014 representing the two years immediately prior to the start of the exchange.

The second interim report[14] provides a more detailed overview of the quasi-experimental research design that was used to statistically examine whether a primary schools participation in the MTE led to greater improvement in pupil level KS1 and KS2 maths attainment compared with none participation. As explained in the second interim report, propensity scores were used to match each of the 47 exchange schools[15] with 20 statistically 'similar' contrast control schools using school-level data for 2014. The purpose of the contrast sample is to capture

[14] See second interim report available at https://www.gov.uk/government/publications/evaluation-of-the-maths-teacher-exchange-china-and-england
[15] As detailed in section 5, 48 schools participated in the Mathematics Teacher Exchange in 2014/15. Of these 48 schools, one was an infant school and not included in the propensity matches because this was based on school level KS2 data, this means that all of the impact analyses relate to 47 MTE schools (44 primary and 3 junior) .

temporal change in KS1/KS2 maths attainment (known technically as the 'counterfactual'). A positive impact would be indicated when the change in attainment observed in the exchange school sample is greater than the change observed in the contrast school sample.

KS1 or KS2 maths attainment of pupils in exchange schools is compared with the attainment of pupils in the contrast schools. Analyses that showed very similar levels of attainment in 2013 and 2014 but an increasing difference 2015 to 2017 would point to evidence that school participation in the exchange led to a positive pupil level impact for KS1 or KS2 maths attainment.

This section is organised into five subsections:

1. School level descriptive analyses 2013 to 2017.
    - A statistical comparison of MTE and contrast control schools using school level KS2 School Census data.
2. Pupil level descriptive analyses 2013 to 2017.
    - A statistical comparison of KS1 and KS2 maths attainment for pupils in MTE schools with pupils in contrast control schools using NPD data.
3. Main (headline) multilevel impact analyses 2013 to 2017.
    - Multilevel (school and pupil) analysis of KS1 and KS2 maths attainment comparing pupils in exchange schools with pupils in contrast control schools.
4. Sensitivity Analyses
    - The main (headline) KS1 and KS2 maths attainment impact analyses are statistically scrutinised for robustness.
5. Scrutinising impact across the separate exchange-contrast school samples
    - Analyses that look at each MTE school and compare attainment for pupils at that school with the attainment of pupils in one of the sample of matched contrast control schools.
6. In addition further sensitivity analysis will be undertaken with regard to high implementation / mastery exploratory analyses
    - Using data gathered in strand one, the relationship between fidelity and KS1 / KS2 maths attainment will be examined
    - The initial descriptive analyses will identify a sub-sample of exchange schools identified to have met high or minimum MTE implementation thresholds, KS1 and KS2 impact analyses will be re-run using these school

The analyses began with the examination of patterns at the school level in terms of KS2 attainment, KS1 attainment for the KS2 pupil cohort, %FSM, %Female and school size. The school level analyses provide the first perspective on KS2 maths attainment differences in exchange schools compared with contrast schools, but do not take any account of within-school (pupil level) attainment variations.

The pupil level descriptive analyses provide the second perspective on KS2 maths attainment differences and first perspective on KS1 attainment differences. Whilst these analyses do directly acknowledge (and examine) variations in attainment at the pupil level, they do not take account of how pupils are clustered into primary schools.

The multilevel analyses acknowledge both school level clustering and within-school pupil level attainment variations and provide the most robust analyses from which to estimate the impact of the exchange on KS1 and KS2 maths attainment.

For the descriptive and main (headline) multilevel impact analyses, a similar approach was taken for dealing with missing data as was taken for the analyses reported in the second interim report. Specifically, a (school-level) listwise deletion approach was adopted. This was done to best ensure that the analyses were undertaken on the same samples of exchange and contrast control schools across the (2013 to 2017) five years. In doing this, schools that did not appear on the school level KS2 census in one of the five years was excluded from the analysis. This brings an additional advantage in terms of internal validity, schools that underwent substantial change during the five years (e.g. became an academy, shut down) will not be included. This helps to ensure that the samples of exchange and matched contrast control schools were consistent and none will have undergone a substantial change in governance structure for the five years of the analyses.

Following the main (headline) multilevel impact analyses, two sensitivity analyses were conducted. First, the listwise deletion of missing values criteria was dropped and all analyses were re-run on the raw KS1 and KS2 pupil samples across the five years. Second, school level KS2 data from 2013 was used to re-match the exchange schools such that the exchange-contrast samples were matched using data from both 2013 and 2014.

The next analyses examined the difference in KS2 maths attainment of pupils in MTE schools and pupils within contrast control schools across the separate exchange-contrast control group samples.

School Level Descriptive Analyses 2013 to 2017

Table 1 in the second interim report presented school-level descriptive statistics for 39 of the 47 exchange schools with complete school level KS1 and KS2 attainment data for 2013, 2014 and 2015. The exchange school statistics are shown alongside comparable statistics for the contrast school sample; 718 of the original 780 matches[16] with complete 2013 to 2015 KS1 and KS2 attainment data. This is known as a listwise deletion approach to missing cases (schools) and is done to best ensure the sample of schools shown across all five years of analyses are the same.

A similar listwise approach was adopted for the final analyses. Across the five years (2013 to 2017), school level KS1 and KS2 attainment data was available for 33 of the 47 exchange schools. Similarly, there are 798 contrast control schools with complete KS1 and KS2 for the five years. However, 242 of these contrast schools were matched to one of the 14 exchange schools without complete KS1/KS2 attainment data. This led to identifying a final (listwise) sample of 556 contrast control schools matched to 33 exchange schools where all schools had complete school-level KS1 and KS2 data. Table 39 below summarises the school-level comparison of the exchange school and matched samples. In Table 39 below, the mean difference between the exchange school and matched samples which are converted into (Cohen's d) effect size statistics.

At the school level, there are some suggestions of positive impact in Table 39 in 2017 relating to the scaled KS2 maths attainment outcome (d=+0.20). This needs to be considered alongside the fact that in 2017, the prior KS1 attainment of KS2 pupils in exchange schools was notably higher than pupils in matched schools (d=+0.33).

These school level analyses are insufficiently robust or sensitive to draw firm conclusions about the impact of MTE on KS2 maths attainment. Specifically, the school level analyses take no account of variation at the pupil level. The analyses that follow will examine the descriptive patterns at the pupil level but these will not take account of prior attainment or the clustering of pupils into schools. The final multilevel impact analyses acknowledge the clustering of pupils at both KS1 and KS2 and, for KS2, control for prior KS1 maths attainment.

[16] Each of the 39 exchange school were matched to 20 contrast schools, resulting in a total sample of (39 x 20) 780 primary schools but 62 schools were dropped because of absent KS1 or KS2 attainment data in 2013, 2014 or 2015.

Table 39: MTE Evaluation: School level descriptive analyses 2013 to 2017

	2013			2014			2015			2016			2017		
	MTE	Match		MTE	Match		MTE	Match		MTE	Match		MTE	Match	
KS1 APS of KS2 pupils	15.8	15.6	+0.11	15.7	15.8	-0.06	15.9	15.7	+0.13	16.0	16.0	+0.02	16.5	16.2	+0.33
KS2 APS*	29.9	29.3	+0.39	30.0	29.9	+0.03	30.1	29.8	+0.28	104.9	104.5	+0.16	106.3	105.8	+0.18
KS1-KS2 Maths Value Added**	101.0	100.4	+0.41	100.8	100.7	+0.07	100.7	100.5	+0.17	-	-	-	-	-	-
KS2 Scaled Maths**	-	-	-	-	-	-	-	-	-	104.5	104.3	+0.09	105.7	105.1	+0.20
%Female	49.0	49.1	-0.03	49.5	49.1	+0.13	49.5	49.3	+0.03	48.8	49.4	-0.07	46.5	48.8	-0.28
%FSM (6 years)	21.0	22.2	-0.07	19.9	22.4	-0.15	23.9	26.7	-0.16	26.2	27.0	-0.04	25.8	27.2	-0.07
Mean School size	381	323	+0.38	387	331	+0.36	395	338	+0.35	401	345	+0.34	409	351	+0.34

* The overall KS2 Average Points Score (APS) attainment measure changed scales in 2016.

** **KS1** to KS2 maths value added score was available for 2013 to 2015.

*** KS2 scaled maths score was available for 2016 and 2017

.

Pupil Level KS1 and KS2 Descriptive Analyses 2013 to 2017

Prior to presenting the pupil level descriptive analyses, some changes in KS1 and KS2 assessments within the evaluation period are discussed. In 2017 the KS1 maths assessment moved from a scale[17] to a categorical[18] measure. This made it impossible to follow the same multilevel linear regression approach presented in the second interim report for our analyses of 2016 and 2017 KS1 data. In response to this, the analyses switch to focus on two categorical outcomes for KS1 maths and were undertaken across the full five years. Specifically, a multilevel logistic regression approach was adopted as summarised in Table 40.

Table 40: Approaches for measuring KS1 maths attainment 2013 to 2017

Analytical approach		2013	2014	2015	2016	2017
Multilevel linear regression	Scale KS1 maths average points score.	✓	✓	✓	n/a	n/a
Multilevel binary logistic regression	Whether Attained expected level or higher in KS1 maths	✓	✓	✓	✓	✓
	Whether exceeded expected level in KS1 maths	✓	✓	✓	✓	✓

The approach to identifying the expected and exceeded KS1 maths thresholds changed during the evaluation period. Between 2013 and 2014, pupils who attained level 2 or higher in KS1 maths are identified as having met or exceeded the expected level and pupils who attained level 3 or higher are identified as having exceeded the expected level. In 2016 and 2017, pupils who are identified as either "working at the expected KS1 standard" or "working at greater depth than the expected KS1 standard" in KS1 maths are identified as having met or exceeded the expected level and pupils who are identified as "working at greater depth than the expected KS1 standard" are identified as having exceeded the expected level.

Whilst it would have been ideal to have had the same (scale) outcome measure across all five years, switching to the binary categorical version does allow the five year KS1 maths impact analyses to be undertaken.

[17] KS1 NPD variable / field name = KS1_MATPOINTS
[18] KS1 NPD variable / field name = KS1_MATH_OUTCOME

Changes to KS2 assessments also took place in 2016; the KS2 maths fine points score[19] (used within the second interim report) was available for 2013 to 2015 but in 2016 a new scale was introduced[20]. This change means that a multilevel linear regression approach remained feasible for all five years. Whilst the change in scale from 2016 is not ideal, this change is likely to affect the exchange and matched pupil samples in a similar way. To provide greater consistency across the five years, the raw KS2 maths test score[21] was included as an outcome variable. Table 41 illustrates the approach taken for analyses of KS2 maths attainment.

Table 41: Approaches for measuring KS2 maths attainment 2013 to 2017

Analytical approach		2013	2014	2015	2016	2017
All multilevel linear regression	KS2 Maths Fine points score	✓	✓	✓	n/a	n/a
	KS2 Maths Scaled Score	n/a	n/a	n/a	✓	✓
	KS2 Maths Raw Score	✓	✓	✓	✓	✓

The KS2 fine points score will be used to provide consistency with the interim report (but this outcome is contained to 2013 to 2015). The new KS2 scaled maths score will be used for 2016 and 2017 to reflect current practice in measuring pupil attainment at KS2. The KS2 raw maths score will be used across all five years to provide temporal consistency for the full evaluation period. All analyses will be drawn on to assess the impact of participation in the MTE programme on KS2 maths attainment.

The second interim report, reported pupil level descriptive analyses for 2014 and 2015 in terms of KS1 and KS2 attainment, %female and %FSM. Table 42 and Table 43 below extends these descriptive analyses to cover the full five year period for the KS1 and KS2 pupil samples respectively. No clear evidence of a difference between the exchange and matched school samples in terms of maths attainment between 2015 and 2017 was observed. In all but one instances, pupils in the exchange school samples attained higher on average in KS1 maths compared with pupils in the matched school samples - but the size of difference is small and in the 'Low' EEF effect size band.

[19] KS2 NPD variable / field name = KS2_MATFINE
[20] KS2 NPD variable / field name = KS2_KS2MATSCORE
[21] KS2 NPD variable / field names = KS2_MATTOTMRK [2013 to 2015] ; KS2_MATMRK [2016 & 2017]

The largest difference between the two samples is seen in 2013, two years prior to the start of the exchange. This might indicate a weakness in matching using just 2014 data and is explored further in the sensitivity analyses reported below.

For KS2, the descriptive analyses show even less evidence of a difference between the exchange and matched school samples in terms of maths attainment between 2015 and 2017. This is shown by the smaller effect sizes reported in Table 43.

Once again, the largest difference between the two samples is seen in 2013, and this is explored further in the sensitivity analyses reported below.

Table 42: MTE Evaluation: KS1 Pupil level descriptive analyses 2013 to 2017

KS1 Average Points Score (Overall and KS1 Maths)

	2013			2014			2015			2016			2017		
	MTE	Match	*d*[1]	MTE	Match	*d*	MTE	Match	*d*	MTE	Match		MTE	Match	
KS1 APS[3]	16.5	16.0	*+0.15*	16.4	16.2	*+0.07*	16.6	16.3	*+0.09*	-	-	-	-	-	-
KS1 Maths APS[4]	16.9	16.5	*+0.14*	16.7	16.6	*+0.04*	16.9	16.7	*+0.06*	-	-	-	-	-	-

KS1 Maths Attainment Thresholds[5]

	2013			2014			2015			2016			2017		
	MTE	Match	*OR*[2]	MTE	Match	*OR*	MTE	Match	*OR*	MTE	Match	*OR*	MTE	Match	*OR*
Expected+	95.2%	93.0%	*1.49*	93.3%	94.2%	*0.86*	94.5%	94.6%	*1.08*	78.1%	76.2%	*1.11*	80.4%	78.9%	*1.10*
Exceeded+	30.3%	24.9%	*1.31*	30.7%	26.6%	*1.22*	31.4	28.0	*1.18*	22.7%	19.9%	*1.18*	25.9%	23.2%	*1.16*

KS1 Pupil demographics

	2013		2014		2015		2016		2017	
	MTE	Match	MTE	Match	MTE	Match	MTE	Match	MTE	Match
% Female	47.3%	48.8%	47.7%	49.0%	49.9%	48.4%	51.2%	48.9%	49.4%	49.1%
% FSM (6)	23.0%	23.0%	22.0%	22.0%	20.0%	21.0%	18.0%	20.0%	17.0%	17.0%

1 - d = Cohens d effect size; **2 -** OR = Odds-Ratios; **3 -** Overall KS1 Average Points Score (APS) available 2013 to 2015; **4 -** KS1 Maths Average Points Score (APS) available 2013 to 2015; **5 -** Thresholds of KS1 maths attainment can be viewed across all five years. These identify when a pupil has reached a standard that is expected at KS2 or not (expected+) and whether a pupil reached a standard that surpassed expectations in KS2 maths (exceeded+). Prior to 2015, a pupil was identified as reaching the expected standard when their KS2 maths attainment was at level 2 or higher (which linked to a KS1 maths APS of 13 points or higher). Similarly, prior to 2015, a pupil was identified as exceeding the expected standard when their KS2 maths attainment was at level 3 or higher (which linked to a KS1 maths APS of 21 points or higher). From 2016 KS1 maths tests became purely categorical (no scale measure available) and the change in methodology is seen to be reflected by the sudden change in statistics observed from 2016. From 2016, pupils who were categorised as 'Working at expected standard' or 'Working at greater depth than expected standard' are classed as expected+ whilst pupils who are categorised as 'Working at greater depth than expected standard' are classed as exceeded+.

Table 43: MTE Evaluation: KS2 Pupil level descriptive analyses 2013 to 2017

KS1 Pupil demographics

	2013			2014			2015			2016			2017		
	MTE	**Match**		**MTE**	**Match**		**MTE**	**Match**		**MTE**	**Match**		**MTE**	**Match**	
KS2 Maths FPS[1]	5.07	4.97	+0.12	5.06	5.07	-0.01	5.08	5.01	+0.08	-	-	-	-	-	-
KS2 Maths SMS[2]	-	-	-	-	-	-	-	-	-	104.6	104.1	+0.07	105.5	105.0	+0.07
KS2 Maths Raw Test Scores[3]:															
Total Score	76.0	74.1	+0.10	75.6	76.2	-0.04	76.3	75.0	+0.07	79.3	77.5	+0.08	78.9	77.6	+0.06
Mental Arithmetic	15.5	15.2	+0.07	15.8	16.0	-0.04	14.9	14.9	+0.01	31.6	31.3	+0.05	32.2	31.8	+0.05
Paper A	29.8	29.2	+0.08	30.3	30.7	-0.05	29.8	29.3	+0.08	24.3	23.6	+0.08	24.7	24.4	+0.05
Paper B	30.7	29.8	+0.12	29.4	29.5	-0.01	31.5	30.8	+0.10	23.3	22.6	+0.08	22.0	21.4	+0.06
KS2 APS[4]	29.8	29.3	+0.10	29.9	29.9	0.00	30.1	29.8	+0.08	-	-	-	-	-	-
mean KS2 score[5]	-	-	-	-	-	-	-	-	-	105.2	104.7	+0.07	104.7	104.1	+0.08
KS1 APS (KS2 cohort)[6]	15.7	15.7	0.00	15.6	15.8	-0.07	15.8	15.8	+0.02	16.1	16.1	+0.02	16.6	16.2	+0.11
KS1 Maths APS (KS2 cohort)[7]	16.1	16.2	-0.02	15.9	16.2	-0.08	16.2	16.2	+0.01	16.4	16.3	+0.02	16.8	16.5	+0.11

KS2 Pupil demographics

	2013		2014		2015		2016		2017	
% Female	48.7%	49.3%	49.9%	49.4%	48.7%	49.2%	48.3%	48.8%	46.5%	48.6%
%FSM (6)	27.0%	25.0%	24.0%	25.0%	26.0%	25.0%	27.0%	25.0%	26.0%	26.0%

1 - KS2 Maths Fine Points Score (FPS) available 2013 to 2015;

2 - KS2 Scaled Maths Score (SMS) available 2016 & 2017;

3 - The KS2 maths raw test scores are available for all five years (2013 to 2017) but there was a change in methodology in 2016 (and this is seen with the sharp change statistics in 2016 and 2017 compared with 2013 to 2015). The greatest change is observed with the mean Mental Arithmetic score which reflects how marks on this paper doubled from 20 points in 2015 to 40 points from 2016. Between 2013 and 2015, there were two written KS2 maths papers (Paper A and Paper B) which were renamed in 2016 to Reasoning 1 and 2 respectively. The marks attributed to these written papers reduced from 40 points in 2015 to 35 points from 2016. The result of the changes increased the total KS2 raw test marks available from 100 in 2015 to 110 from 2016. Prior to 2016, the 100 points were weighted 20 / 40 / 40 for arithmetic / paper A / paper B and from 2016 the 110 points were weighted 40 / 35 / 35 for arithmetic / reasoning 1 / reasoning 2.

4 - Overall KS2 Average Points Score (APS) available 2013 to 2015.

5 - Mean KS2 Score - derived from mean score of all (scaled) KS2 test scores - available 2016 & 2017.

6 - Overall KS1 Average Points Score (APS) for KS2 pupil cohort (i.e. for the 2013 KS2 cohort, the KS1 data stems from 2009 when this cohort sat their KS1 tests) - available for all years 2013 to 2017.

7 - KS1 Maths Points Score (APS) for KS2 pupil cohort - available for all years 2013 to 2017.

Main (headline) multilevel Impact analyses

The descriptive analyses showed little / no evidence that participation in the MTE programme resulted in gains in pupil level KS1 or KS2 maths attainment but the multilevel analyses provide a more comprehensive impact assessment. This is because the KS1 and KS2 multilevel analyses will statistically take account of how pupils are clustered into schools. Additionally, the KS2 analyses statistically controls for prior KS1 maths attainment at the pupil level.

Impact at Key Stage 1

For KS1 maths attainment, KS1 Maths Points Score (2013 to 2015) and KS1 expected thresholds (2013 to 2017) are modelled including a single school level binary variable that identified whether a pupil was located in an exchange school (=1) or not (=0). This is known as an outcome-only model.

Table 44 reports the main impact analyses for KS1 maths attainment. For the 2013 to 2014 KS1 maths points score models, Table 45 shows the estimated model coefficient and standard error for the exchange school pupil sample. The coefficient is then converted into Cohen's d effect size statistics with 95% confidence intervals.

For the 2013 to 2017 KS1 maths attainment threshold models, Table 44 shows the estimated model coefficient and standard error for the exchange school pupil sample. The coefficient is then converted into odds-ratio statistics[22] with 95% confidence intervals. The odds-ratios and confidence intervals are then converted into Cohens d effect size estimates using the formula set out by Sanchez-Meca et al (2003). Finally, Table 44 shows the number of primary schools and pupils included into the KS1 maths analyses.

The analyses found that whilst pupils in the exchange schools were more likely to meet or exceed the expected KS1 maths thresholds compared with pupils in the matched contrast school sample, the difference was very small and not statistically significant between 2015 and 2017. This leads us to conclude that from the main impact analyses, we found no evidence that a schools participation in the Shanghai mathematics teacher exchange resulted in gains in pupil attainment in KS1 maths.

The largest difference shown in Table 44 was in 2013. This echoes the descriptive finding shown in Table 43 and may be an indication of a weakness in matching using just 2014 data and this is explored further in the sensitivity analyses reported below.

[22] Relative odds of pupils in exchange schools reaching the threshold compared with pupils in the matched schools.

Table 44: Pupil Level KS1 Attainment Models

KS1 Maths Points Score (2013 to 2015)

Multilevel Linear regression Analyses

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2013	0.44*	0.200	+0.13*	+0.01; +0.24	551	24,509
2014	0.09	0.191	+0.03	-0.08; +0.14	552	25,435
2015	0.12	0.186	+0.04	-0.07; +0.14	553	26,348

KS1 Maths Attainment Thresholds (2013 to 2017)

Multilevel Logistic regression Analyses

Exceeding expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2013	0.27*	0.128	1.31*	1.02; 1.69	+0.15*	+0.01; +0.29	551	24,509
2014	0.15	0.116	1.16	0.93; 1.46	+0.08	-0.04; +0.21	552	25,435
2015	0.13	0.111	1.14	0.92; 1.42	+0.07	-0.05; +0.19	553	26,348
2016	0.14	0.144	1.15	0.87; 1.53	+0.08	-0.08; +0.23	553	26,028
2017	0.16	0.113	1.17	0.94; 1.46	+0.09	-0.03; +0.21	540	26,010

Meeting expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2013	0.36*	0.184	1.44*	1.00; 2.06	+0.20*	0.00; +0.40	551	24,509
2014	-0.05	0.177	0.95	0.67; 1.34	-0.03	-0.22; +0.16	552	25,435
2015	0.00	0.189	1.00	0.69; 1.45	0.00	-0.20; +0.21	553	26,348
2016	0.10	0.111	1.10	0.88; 1.37	+0.05	-0.07; +0.17	553	26,028
2017	0.14	0.095	1.15	0.96; 1.39	+0.08	-0.02; +0.18	540	26,010

* $p<0.05$

1 - converting Odds Ratio to Cohens d effect size

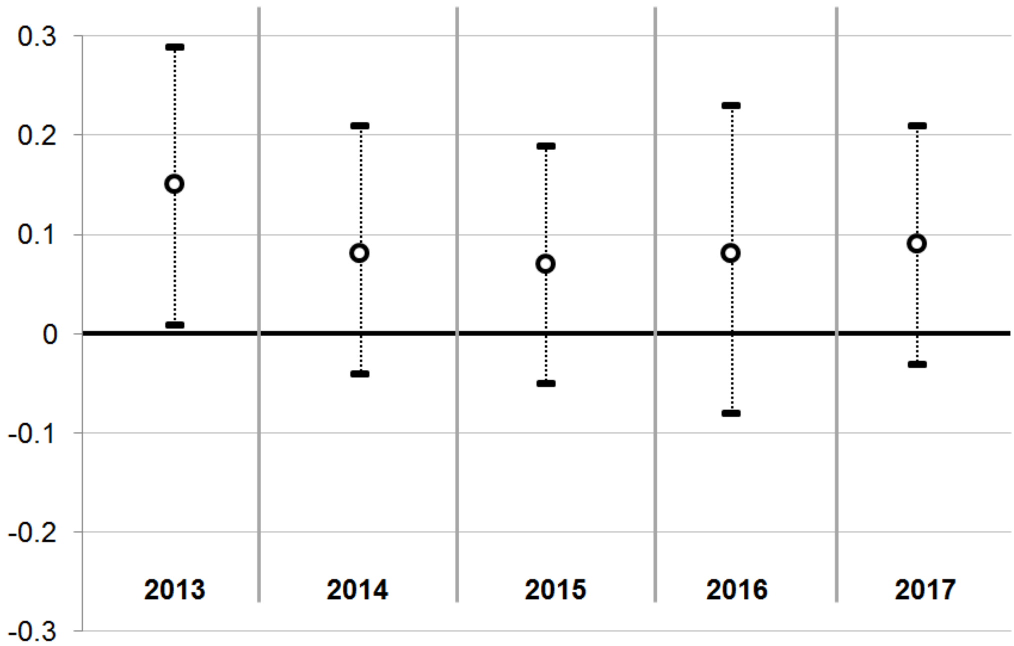
Figure 6 uses error bars to illustrate the Cohen's d effect size statistics from the multilevel models using the 'exceeding expected KS1 maths level' outcome variable. The estimated effect size is shown as the circle in the centre of an upper and lower bar which shows the 95% confidence intervals from the multilevel analyses.

As can be seen, a small positive effect is seen across all five years but this is not statistically significantly different from zero in 2014 to 2017. Further, the chart shows no evidence of an increasing difference between the exchange and matched samples between 2015 and 2017.

Figure 6: KS1 maths attainment difference between pupils in MTE schools and pupils in contrast control schools 2013 to 2017

Coefficient from multilevel logistic analyses, odds-ratios converted into Cohen's d effect size statistics

Outcome: *Whether a pupil exceeds expected KS1 maths level (=1) or not (=0)*



Impact at Key Stage 2

For KS2 maths attainment, KS2 Maths Fine Points Score (2013 to 2015); KS2 scaled maths score (2016 to 2017) and KS2 Maths Raw Score (2013 to 2016) are modelled. The models are constructed in two stages; First, an outcome-only model and then including KS1 maths attainment as an explanatory variable (creating a KS1 to KS2 value added model).

Table 45 reports the main impact analyses for KS2 maths attainment. For all models Table 45 shows the estimated model coefficient and standard error for the exchange school pupil sample. The coefficient is then converted into Cohen's d effect size statistics with 95% confidence intervals.

The analyses found that the KS2 maths attainment for pupils in the exchange schools was comparable with pupils in the matched contrast school sample, no statistically significant difference was observed between 2015 and 2017. This leads us to conclude that from the main impact analyses, we found no evidence that a schools participation in the Shanghai mathematics teacher exchange programme resulted in gains in pupil attainment in KS2 maths.

The largest difference shown in Table 45 was in 2013. This echoes the descriptive finding shown in Table 45 and may be an indication of a weakness in matching using just 2014 data and this is explored further in the sensitivity analyses reported below.

Table 45: Pupil Level KS2 Attainment Models Maths Fine Point Score (2013 to 2015)

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2013						
Outcome Only	0.12*	0.051	+0.14*	+0.02; +0.26	589	24,810
Value Added	0.11*	0.046	+0.13*	+0.02; +0.24	589	23,755
2014						
Outcome Only	0.01	0.048	+0.01	-0.11; +0.12	589	25,746
Value Added	0.01	0.040	+0.02	-0.08; +0.11	589	24,646
2015						
Outcome Only	0.07	0.049	+0.09	-0.03; +0.20	589	26,260
Value Added	0.04	0.042	+0.05	-0.05; +0.15	589	25,121

Scaled Maths Score (2016 to 2017)

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2016						
Outcome Only	0.25	0.470	+0.03	-0.09; +0.16	589	26,654
Value Added	0.13	0.420	+0.02	-0.09; +0.13	589	25,555

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2017						
Outcome Only	0.56	0.495	+0.07	-0.05; +0.19	589	27,506
Value Added	-0.13	0.437	-0.02	-0.12; +0.09	589	26,262

Maths Raw Point Score (2013 to 2017)

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2013						
Outcome Only	2.31	1.209	+0.12	0.00; +0.24	589	24,498
Value Added	2.20	1.135	+0.11	0.00; +0.23	589	23,518
2014						
Outcome Only	-0.35	1.036	-0.02	-0.13; +0.09	589	25,455
Value Added	-0.14	0.923	-0.01	-0.11; +0.09	589	24,439
2015						
Outcome Only	1.45	1.096	+0.08	-0.04; +0.20	589	25,937
Value Added	0.74	1.013	+0.04	-0.07; +0.15	589	24,884
2016						
Outcome Only	0.90	1.533	+0.04	-0.09; +0.17	589	26,363
Value Added	0.44	1.422	+0.02	-0.10; +0.14	589	25,311
2017						
Outcome Only	1.50	1.583	+0.06	-0.07; +0.20	589	27,196
Value Added	-0.33	1.434	-0.01	-0.13; +0.11	589	26,022

* $p<0.05$

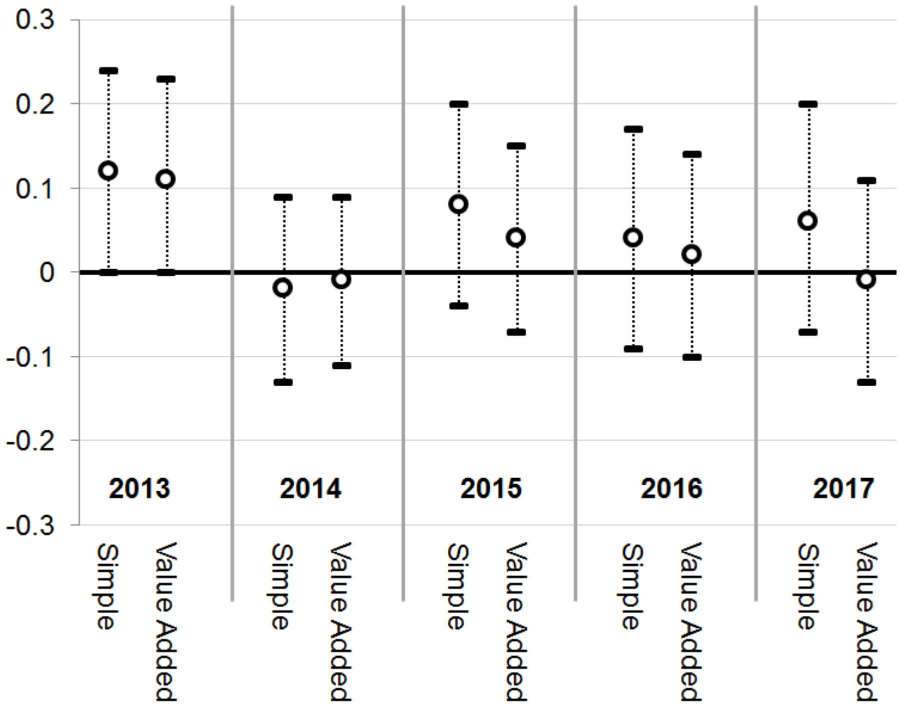
Figure 7 uses error bars to illustrate the Cohen's d effect size statistics from the multilevel models using the raw KS2 maths test outcome variable. The estimated effect size is shown as the circle in the centre of an upper and lower bar which shows 95% confidence intervals from the multilevel analyses.

As can be see, the effect size remains close to zero 2014 to 2017. Further, once prior KS1 maths attainment is controlled for, the effect size moves even closer towards zero between 2015 and 2017.

Figure 7: KS2 maths attainment difference between pupils in MTE schools and pupils in contrast control schools 2013 to 2017

Coefficient from multilevel analyses converted into Cohen's d effect size statistics with 95% confidence intervals

Outcome: *Raw KS2 maths test score*



Sensitivity analyses

As outlined above, the descriptive and main (headline) impact analyses were undertaken on a sub-sample of exchange and matched contrast control schools, restricted using a school-level listwise deletion approach to missing values. This led to a sample of 33 exchange schools matched to 556 contrast schools with school level KS1 and KS2 attainment data for the KS2 pupil cohort for the full (2013 to 2017) five year period. This approach best ensured that the same sample of schools were included in the analyses across the five years.

For sensitivity analyses, all of the impact analyses were re-run on two different samples:

The Raw sample: The size of the exchange and contrast control school sample will fluctuate over the five year period with a maximum size of all 47 exchange schools matched to 940 contrast control schools.

Re-matched Sample: Sample of exchange to matched schools were re-matched using 2013 school level data. This was done to ensure that the matched sample reflected the exchange school sample across 2013 and 2014 (the two years prior to the exchange) rather than just 2014 (used for the main analyses). Re-matching resulted in a sample of 27 exchange schools matched to 179 contrast schools.

Table 46 summarises the KS1 models for the raw and re-matched school samples. Specifically, Table 46 summarises the multilevel logistic models for the higher KS1 threshold outcome (whether a pupil exceeded the expected level in KS1 maths or not).

For KS1 maths, both sensitivity analyses echo the patterns observed in the main impact analyses. These findings re-affirm our conclusion that we found no evidence that participation in the exchange resulted in gains in KS1 maths.

Table 47 summarises the KS2 models for the raw and re-matched school samples. Specifically, Table 47 summarises the multilevel linear regression models for the raw KS2 maths test score outcome.

For KS2 maths, both sensitivity analyses echo the patterns observed in the main impact analyses. These findings also re-affirm our conclusion that we found no evidence that participation in the exchange resulted in gains in KS2 maths.

Table 46: KS1 Maths Attainment Sensitivity Analyses

Exceeding expected KS1 level in maths

Raw Sample

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2013	0.26*	0.114	1.30*	1.04; 1.63	+0.15	+0.02; +0.27	906	39,952
2014	0.12	0.100	1.12	0.92; 1.36	+0.06	-0.04; +0.17	902	41,412
2015	0.14	0.095	1.15	0.95; 1.38	+0.07	-0.03; +0.18	874	41,578
2016	0.24	0.136	1.28	0.98; 1.67	+0.13	-0.01; +0.28	842	39,351
2017	0.19	0.107	1.21	0.98; 1.49	+0.10	-0.01; +0.22	800	38,364

Rematched sample

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2013	0.14	0.161	1.15	0.84; 1.58	+0.08	-0.09; +0.25	192	9,391
2014	0.03	0.146	1.03	0.78; 1.38	+0.02	-0.14; +0.18	193	9,751
2015	0.01	0.140	1.01	0.77; 1.33	+0.01	-0.14; +0.16	193	9,951
2016	0.04	0.178	1.04	0.73; 1.30	+0.02	-0.17; +0.21	193	9,945
2017	0.01	0.132	1.01	0.78; 1.30	0.00	-0.14; +0.15	187	9,939

* p<0.05

Table 47: KS2 Maths Attainment Sensitivity Analyses

Maths Raw Point Score

Raw Sample

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2013						
Outcome Only	2.79*	1.182	+0.14*	+0.02; +0.26	932	37,989
Value Added	2.30*	1.077	+0.12*	+0.01; +0.22	932	36,444
2014						
Outcome Only	0.11	1.044	+0.01	-0.10; +0.12	986	42,310
Value Added	-0.28	0.878	-0.01	-0.11; +0.08	986	40,571
2015						
Outcome Only	1.86	1.014	+0.10	-0.01; +0.21	959	42,096
Value Added	1.52	0.879	+0.08	-0.01; +0.18	959	40,393
2016						
Outcome Only	1.42	1.402	+0.06	-0.06; +0.18	922	41,141
Value Added	0.92	1.248	+0.04	-0.07; +0.15	922	39,483
2017						
Outcome Only	1.59	1.473	+0.07	-0.05; +0.19	884	40,607
Value Added	-0.06	1.308	0.00	-0.11; +0.11	884	38,863

Re-matched Sample

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2013						
Outcome Only	-0.08	1.623	0.00	-0.17; +0.16	206	9,343
Value Added	1.29	1.392	+0.07	-0.07; +0.21	206	8,930
2014						
Outcome Only	-1.36	1.437	-0.08	-0.23; +0.08	206	9,799
Value Added	0.26	1.193	+0.01	-0.12; +0.14	206	9,399

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2015						
Outcome Only	1.24	1.466	+0.07	-0.09; +0.23	206	9,882
Value Added	1.66	1.196	+0.09	-0.04; +0.22	206	9,463
2016						
Outcome Only	-0.61	1.834	-0.03	-0.19; +0.13	206	10,162
Value Added	0.46	1.515	+0.02	-0.11; +0.15	206	9,741
2017						
Outcome Only	0.82	1.944	+0.03	-0.13; +0.20	206	10,548
Value Added	0.17	1.701	+0.01	-0.13; +0.15	206	10,053

* $p<0.05$

Scrutinising impact across the separate exchange-contrast school samples

The analyses presented prior to this section focused on comparing the KS1 and KS2 maths attainment for pupils located in a sample 47 MTE schools with the attainment for pupils located in the sample of matched contrast control schools. Impact at KS1 and KS2 was examined across all MTE schools, and these analyses found no evidence that participation in the exchange led to pupil gains in KS1 or KS2 maths attainment.

In this section we look closer at the separate exchange-contrast control school samples. These exploratory analyses focus solely on KS2 maths attainment. The shift towards measuring maths attainment using very blunt thresholds in 2016 limits the scope and value of follow on exploratory analyses for KS1 maths. Raw KS2 maths test scores are examined amongst the listwise sample of 33 exchange schools and 556 contrast control schools. For each of the 33 exchange schools, the mean KS2 maths attainment is compared with the 33 matched samples of contrast control schools from 2013 to 2017.

Table 48 below summarises these analyses by first indicating the number of instances when pupil KS2 maths attainment within an exchange school was greater (by an effect size of d=+0.02 or greater) than the pupil attainment of the matched contrast control sample. Table 48 also indicates the number of instances when pupil KS2 maths attainment within an exchange school was lower (by an effect size of d= -0.02 or lower) than the pupil attainment of the matched contrast control sample. Table 48 finally provides additional detail on the range of positive and negative effect sizes found across the 33 exchange-matched school sub-samples.

Table 48: Comparing KS2 maths attainment for pupils in 33 exchange schools with their 33 matched contrast control school samples.

	2013	2014	2015	2016	2017
Exchange > Matched Sample	21 (64%)	14 (42%)	23 (70%)	19 (58%)	18 (55%)
Exchange = Matched Sample	3 (9%)	5 (15%)	2 (6%)	1 (3%)	2 (6%)
Exchange < Matched Sample	9 (27%)	14 (42%)	8 (24%)	13 (39%)	13 (39%)
Cohen's d effect sizes for +VE Impact [Exchange > Matched Sample]					
High/V High (+0.45 or higher)	4 (12%)	0 (0%)	1 (3%)	2 (6%)	5 (15%)
Moderate (+0.19 to <+0.45)	9 (27%)	2 (6%)	13 (39%)	8 (24%)	7 (21%)
Low (+0.02 to <+0.19)	8 (24%)	12 (36%)	9 (27%)	9 (27%)	6 (18%)
Cohen's d effect sizes for -VE Impact [Exchange < Matched Sample]					
Low (+0.02 to <+0.19)	4 (12%)	14 (42%)	3 (9%)	5 (15%)	5 (15%)
Moderate (-0.19 to -0.45)	5 (15%)	0 (0%)	2 (6%)	7 (21%)	6 (18%)
High/V High (below -0.45)	0 (0%)	0 (0%)	3 (9%)	1 (3%)	2 (6%)

Table 48 shows a clear balance between the exchange school and contrast control school samples in 2014. In 2014, in 14 of the 33 exchange-contrast sub-samples, mean pupil KS2 maths attainment was higher in the exchange school sample (in two sub-samples, this equated to an effect size above +0.19 whilst for the remaining 12 it equated to an effect size between +0.02 and +0.19). At the same time, in 2014 there are 14 of the 33 exchange-contrast sub-samples where mean pupil KS2 maths attainment was lower in the exchange school sample (in all 14 sub-samples, the difference equated to an effect size between -0.02 and -0.19). Essentially, the exchange-contrast sub-samples are fairly tightly clustered either side of a zero effect size. This balance is good to see, and was expected given that 2014 was the year in which school level data was used to match the exchange schools with a sample of statistically similar contrast control schools.

Between 2015 and 2017, Table 48 shows the effect sizes across the 33 exchange-contrast sub-samples diverging - but this was in both directions.

For example, in 2017, in 18 of the 33 exchange-contrast sub-samples, mean pupil KS2 maths attainment was higher in the exchange school sample. Additionally, greater variation in the positive impact was observed in 2017 compared with 2014. Specifically, in five sub-samples, it equated to an effect size of +0.45 sds or higher; in seven sub-samples it equated to an effect size between +0.19 to less than +0.45, and for the remaining six sub-samples it equated to an effect size between +0.02 and +0.19. At the same time, in 2017 there are 13 of the 33 exchange-contrast sub-samples where mean pupil KS2 maths attainment was lower in the exchange school sample. Once again, greater variation in the negative impact is observed in these 2017 13 exchange-contrast sub-samples compared with 2014. Specifically, in two sub-samples, it equated to an effect size of -0.45 sds or lower; in six sub-samples it equated to an effect size between -0.19 to greater than -0.45, and for the remaining five sub-samples it equated to an effect size between -0.02 to greater than -0.19.

In summary, looking across the separate exchange-contrast school sub-samples does not result in finding evidence that participation in the exchange led to gains in pupil KS2 maths attainment. Across the 33 sub-samples, the instances where mean attainment is greater amongst pupils in exchange school samples are offset by other instances where mean attainment is greater amongst pupils in the matched contrast control school samples.

In summary, from the analyses to this point we have found no evidence to suggest that a primary schools participation in the Shanghai mathematics teacher exchange programme led to gains in pupil attainment in KS1 or KS2 maths.

High implementation / mastery exploratory analyses

Implementation / mastery data for 2016 and 2017 was obtained from 37 of the 47 MTE schools (35 primary and 2 junior). Sub-samples of schools with high implementation / mastery at KS1 and KS2 were identified using the criteria set out in section 8. Follow-on statistical analyses will focus on the sub-sample of high implementation / mastery schools. These analyses will examine if/how KS1 and KS2 maths attainment of pupils within the high implementation / mastery MTE sub-sample of schools differs from the maths attainment of pupils in matched contrast schools. Given that schools with high implementation / mastery at KS1 and/or KS2 were identified using data from 2016 and 2017, the follow-on impact analyses will focus only on KS1 and KS2 maths attainment in 2017. KS1 and KS2 maths attainment in 2014 will also be examined in order to provide a baseline.

Table 49 provides a numerical summary of schools and pupils in the MTE and contrast samples for the KS1 and KS2. A listwise deletion approach has been taken for schools and pupils with data missing in either 2014 or 2017. At KS1, 12 schools are identified as having high implementation and mastery and with KS1 maths data in both 2014 and 2017. These 12 high mastery & implementation schools are matched to 208 schools in 2014 and 205 schools in 2017[23].

Table 49: Numbers of schools & pupils in high implementation/mastery analyses 2014 & 2017

	Mastery & High Implementation in KS1							
	2014				2017			
	MTE		Contrast		MTE		Contrast	
	Schools	Pupils	Schools	Pupils	Schools	Pupils	Schools	Pupils
High Mastery	12	671	208	9,574	12	659	205	9,916
Mastery+	16	925	271	12,396	16	913	265	12,870

	Mastery & High Implementation in KS2							
	2014				2017			
	MTE		Contrast		MTE		Contrast	
	Schools	Pupils	Schools	Pupils	Schools	Pupils	Schools	Pupils
High Mastery	9	418	153	6,530	9	437	153	6,929
Mastery+	10	450	171	7,169	10	467	171	7,614

Table 50 presents some descriptive statistics comparing the MTE and matched contrast samples in 2014 and 2017. For both KS1 maths attainment thresholds, a gap is observed to open up between the MTE and contrast samples between 2014 and 2017. These analyses provide the first suggestion that participation in the MTE exchange led to pupil gains in maths attainment.

In 2014, pupils in MTE schools with high KS1 implementation are seen to be equally as likely to reach the expected KS1 level as pupils in matched contrast schools but by 2017, pupils in MTE schools with high KS1 implementation are seen to be more likely to reach the expected KS1 level as pupils in matched contrast schools. The difference is greater amongst pupils in MTE schools with both high mastery and high KS1 implementation; in 2017 these pupils are 1.49 times as likely to attain the expected KS1 level in maths compared with pupils in matched contrast schools. Pupils in mastery+ schools with high KS1 implementation are observed to be 1.39

[23] The slight (208 and 205) fluctuation is because KS2 data was used in the matching and some of these matches may be junior schools with no KS1 data. The numbers of schools in 2014 and 2017 for the KS2 analyses are exactly the same, reflecting the use of KS2 data in the match.

times as likely as pupils in matched contrast schools to attain the expected KS1 level in maths.

In 2014, pupils in MTE schools with high KS1 implementation are seen to be more likely to exceed the expected KS1 level as pupils in matched contrast schools and this difference is seen to widen by 2017. Again, the difference is greater amongst pupils in MTE schools with both high mastery and high KS1 implementation. In 2014, these pupils are observed to be 1.42 times as likely to exceed the expected KS1 level in maths compared with pupils in matched contrast schools. In 2017 these pupils are observed to be 1.72 times as likely to exceed the expected KS1 level in maths compared with pupils in matched contrast schools. Pupils in mastery+ schools with high KS1 implementation are observed to be 1.34 times as likely as pupils in matched contrast schools to exceed the expected KS1 level in maths in 2014 which is seen to increase to being 1.46 times as likely in 2017.

Table 50 also shows that the MTE and matched contrast sample remained relatively balanced in terms of gender. However, whilst the proportion of disadvantaged pupils in MTE schools is comparable with the proportion of disadvantaged pupils in matched contrast schools in 2014, a difference is evident in 2017. Specifically, in 2017 pupils in matched contrast schools were more likely to be classed as FSM compared with pupils in MTE schools.

Whilst these descriptive analyses do provide the first evidence that participation in the MTE led to pupil gains in KS1 maths attainment, some caution is advised. First, these are bivariate descriptive analyses that do not take account of the clustering of pupils into schools; second, the observed %FSM imbalance in 2017 and third, KS1 maths attainment is a teacher assessment that uses broad/course attainment categories. The first two of these will be addressed within the multilevel logistic regression analyses that directly acknowledge the clustering of pupils into schools and will allow the inclusion of a pupil level FSM covariate to address the observed 2017 imbalance. The second two cannot be statistically addressed but do need to be kept in mind when interpreting impact at KS1.

Table 50: MTE Evaluation: KS1 Pupil level descriptive analyses 2013 to 2017

KS1 Maths Attainment Thresholds

	2014			2017		
	MTE	**Match**	***OR***	**MTE**	**Match**	***OR***
High mastery & KS1 implementation MTE schools						
Expected+	94.5%	94.5%	1.00	85.1%	79.3%	1.49
Exceeded+	34.9%	27.4%	1.42	34.3%	23.3%	1.72
Mastery+ & high KS1 implementation MTE schools						
Expected+	94.2%	94.2%	1.00	83.9%	78.9%	1.39
Exceeded+	32.6%	26.5%	1.34	30.3%	22.9%	1.46

KS1 Pupil demographics

	2014		2017	
	MTE	**Match**	**MTE**	**Match**
High mastery & KS1 implementation MTE schools				
% Female	47.7%	48.7%	48.3%	48.4%
% FSM (6)	20.6%	20.6%	13.5%	17.1%
Mastery+ & high KS1 implementation MTE schools				
% Female	47.5%	48.8%	47.3%	48.7%
% FSM (6)	21.5%	22.4%	13.8%	18.1%

OR = Odds-Ratios

Table 51 presents the KS1 maths multilevel logistic analyses. Two KS1 maths attainment thresholds are shown; exceeding the expected level and meeting the expected level. **Please note** that these KS1 threshold analyses relate solely to the subsample of 12 MTE schools with High Mastery and KS1 implementation and/or the subsample of 16 MTE schools with Mastery+ and high KS1 implementation.

In 2014, no statistically significant difference is observed between the MTE and matched contrast samples. This is seen with both the high mastery & implementation sub-sample and with the mastery and high implementation sub-sample.

In 2017, pupils in MTE schools are observed to be statistically significantly more likely to attain the KS1 maths attainment thresholds compared with pupils in the matched contrast school sample. The difference is greater at the higher threshold and remains positive and statistically significant when the FSM covariate is included into the model (not shown in Table 51).

These high implementation and mastery analyses have revealed evidence that participation on the Shanghai exchange led to positive gains in KS1 maths attainment for some MTE schools. It seems that the level of implementation and mastery are important factors in determining impact on pupil KS1 maths attainment.

Table 51: KS1 Maths attainment models 2014 & 2017

High Mastery & KS1 Implementation

Exceeding expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2014	0.24	0.182	1.27	0.89; 1.81	+0.13	-0.07; +0.33	220	10,235
2017	0.57*	0.188	1.77*	1.22; 2.55	+0.31*	+0.11; +0.52	217	10,562

Meeting expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2014	0.04	0.288	1.04	0.59; 1.83	+0.02	-0.29; +0.33	220	10,235
2017	0.40*	0.169	1.50*	1.08; 2.08	+0.22	+0.04; +0.40	217	10,562

Mastery+ & High KS1 Implementation

Exceeding expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2014	0.21	0.158	1.24	0.91; 1.69	+0.12	-0.05; +0.29	287	13,306
2017	0.42*	0.159	1.53*	1.12; 2.09	+0.23*	+0.06; +0.41	281	13,766

Meeting expected KS1 level in maths

	Coef	s.e.	OR	95% CIs for OR	d[1]	95% CIs for d	n schools	n pupils
2014	-0.01	0.238	0.99	0.62; 1.57	-0.01	-0.27; +0.25	287	13,306
2017	0.34*	0.139	1.41*	1.07; 1.85	+0.19	+0.04; +0.34	281	13,766

* $p<0.05$

1 - converting Odds Ratio to Cohens d effect size

The next analyses consider impact at KS2 amongst high implementation and mastery MTE schools. Table 52 presents descriptive statistics comparing MTE and contrast school samples in terms of maths attainment, gender and %FSM. As with KS1, MTE schools have a high level of implementation at KS2 and two levels of mastery are shown (high mastery and mastery+).

In 2014, the KS2 maths attainment of pupils in MTE schools with high implementation was slightly lower than pupils in matched contrast schools but in 2017 this pattern is seen to reverse.

For schools with both high implementation and mastery, in terms of the mean overall KS2 maths score, a negative effect size is observed in 2014 (d=-0.07) but this changes to a slightly larger positive effect size by 2017 (d=+0.13). In terms of pupil demographics, the MTE and matched sample seem reasonably comparable in terms of gender but there is a slightly larger proportion of disadvantaged pupils in MTE schools compared with the matched contrast school sample.

The multilevel linear regression analyses provide a more comprehensive assessment of the KS2 maths attainment differences between the MTE and contrast school samples. Specifically, these analyses will acknowledge the clustering of pupils into schools and statistically take account of (or control for) other explanatory variables (such as KS1 maths attainment and FSM status).

Table 52: MTE Evaluation - high KS2 implementation & mastery descriptive analyses

KS2 Attainment (Overall and KS1 Maths)

	2014			2017		
	MTE	**Match**	**d**	**MTE**	**Match**	**d**
High Mastery & KS2 implementation						
Total Score	75.8	77.1	-0.07	81.6	78.5	+0.13
Mental Arithmetic	15.6	16.3	-0.18	32.6	32.0	+0.08
Paper A	30.4	31.0	-0.08	26.1	24.7	+0.16
Paper B	29.9	29.8	+0.01	22.9	21.8	+0.12
KS1 Maths APS (KS2 cohort)	15.7	16.2	-0.15	16.7	16.4	+0.08
Mastery+ & high KS2 implementation						
Total Score	75.4	76.2	-0.07	82.0	78.2	+0.16
Mental Arithmetic	15.5	16.2	-0.17	32.7	31.9	+0.11
Paper A	30.2	30.9	-0.09	26.2	24.6	+0.18
Paper B	29.7	29.6	+0.01	23.1	21.7	+0.16
KS1 Maths APS (KS2 cohort)	15.7	16.2	-0.12	16.7	16.3	+0.09

KS2 Pupil demographics

	2014		2017	
High Mastery & KS2 implementation				
% Female	46.9%	49.8%	45.8%	49.7%
%FSM (6)	24.8%	23.8%	29.7%	24.6%
Mastery+ & high KS2 implementation				
% Female	46.9%	49.5%	45.4%	49.7%
%FSM (6)	24.3%	24.0%	28.7%	24.9%

Table 53 presents the KS2 maths multilevel linear analyses. The models presented are for the overall KS2 maths test score in 2014 and 2017. As with previous impact analyses, an outcome only and value added model are shown. The outcome only models just included a school level explanatory variable that identified whether a pupil was located in an MTE (=1) or contrast school (=0). The value added models include an additional pupil level KS1 maths attainment explanatory variable and therefore statistically control for pupil level variations in prior maths attainment. Following the slight FSM imbalance observed from the descriptive analyses (Table 52), a further model stage included pupil level FSM status but these are not shown in Table 53; any impact of including the FSM variable is noted in the text. All of these models were replicated across the three separate KS2 maths papers and any findings are discussed in the text.

In 2014, no statistically significant difference is observed between the MTE and matched contrast samples. This is seen with both the high mastery & implementation sub-sample and with the mastery and high implementation sub-sample.

In 2017, no statistically significant difference is observed between the MTE and matched contrast samples. This is seen with both the high mastery & implementation sub-sample and with the mastery and high implementation sub-sample. The same pattern of no statistically significant difference was found across the three KS2 maths papers and when FSM was included into the models (not shown in Table 53).

In summary, whilst there is descriptive evidence that KS2 maths attainment was higher in high implementation / mastery schools compared with the attainment of pupils in matched contrast schools, the multilevel analyses reveal that this difference is not statistically significant.

Table 53: MTE Evaluation: KS2 Pupil level descriptive analyses 2014 & 2017

Maths Raw Point Score

High Mastery & KS2 Implementation

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2014						
Outcome Only	0.12	2.309	+0.01	-0.25; +0.26	162	6,850
Value Added	0.81	1.971	+0.05	-0.17; +0.26	162	6,561
2017						
Outcome Only	3.73	3.551	+0.16	-0.14; +0.45	162	7,257
Value Added	1.57	2.900	+0.07	-0.17; +0.30	162	6,900

Mastery+ & KS2 Implementation

	Coef	s.e.	d	95% CIs for d	n schools	n pupils
2014						
Outcome Only	-0.20	2.119	-0.01	-0.24; +0.22	181	7,513
Value Added	0.23	1.817	+0.01	-0.19; +0.21	181	7,208
2017						
Outcome Only	4.65	3.281	+0.20	-0.08; +0.47	181	7,964
Value Added	2.43	2.697	+0.10	-0.12; +0.32	181	7,588

No statistically significant difference found across the three KS2 maths papers

Re-matching using 2013 school level KS2 data

As outlined in the second interim report, in response to the observed school level imbalance in 2013 (two years prior to the start of the MTE), the sample of 47 MTE and 940 matched contrast control schools were re-matched so that the matching process drew on school level data from both 2014 (as original) and 2013.

Using 2013 school level KS2 data, the original 47:940 school sample was re-matched. This was done by first creating a school-level binary outcome for 2013 for the KS2 APS score (1= above median of 29.2; 0=median or lower). The model included the following explanatory variables; KS1 attainment, School size; Maths KS1-KS2 VA Score; %FSM; %Female; %SEN. Propensity scores were then generated and within each of the 47 sub-samples, the matched schools were rank ordered according to how closely the propensity scores for the contrast schools matched their respective exchange school.

This resulted in a reduction in the exchange school sample from 47 to 33 schools, because 14 exchange schools did not have 2013 data. These 14 exchange schools and their 280 matched contrast schools were dropped.

This reduced the sample to 33 schools with 660 matches. The propensity scores were then used to exclude contrast schools that either did not have 2013 data or where the 2013 propensity scores did not match closely with the exchange school they were matched to using 2014 data. This process reduced the contrast control sample down to 218 schools; the 33:218 re-matched sub-sample.

Following the re-matching process, listwise deletion of missing cases was re-done to ensure that all exchange and control schools are matched using 2013 and 2014 data AND to ensure that all these schools had complete school level KS2 attainment data for 2013 to 2017 inclusive. This reduced the sample of exchange schools further to 27 and the sample of matched contrast schools to 179.

Re-matched sample:
27 exchange schools matched to 179 contrast schools with complete KS1 and KS2 data for 2013 to 2017 inclusive.

Table 54 reconstructs the school level descriptive analyses shown in Table 2 of the second interim report but using the 27:179 re-matched sample.

Table 55 reconstructs the pupil level KS1 descriptive analyses shown in Table 3 of second interim report but using the 27:179 re-matched sample.

Table 56 reconstructs the pupil level KS2 descriptive analyses shown in Table 3 of second interim report but using the 27:179 re-matched sample.

Table 54: School level descriptive analyses 2013 to 2017 [re-matched sample]

	2013			2014			2015			2016			2017		
	MTE	Match		MTE	Match		MTE	Match		MTE	Match		MTE	Match	
KS1 APS of KS2 pupils	15.6	15.9	-0.19	15.5	15.9	-0.30	15.7	15.9	-0.10	15.9	16.2	-0.27	16.5	16.3	+0.12
KS2 APS*	29.8	29.7	+0.04	29.9	30.2	-0.19	30.2	30.0	+0.12	105.0	105.1	-0.01	106.4	106.2	+0.07
KS1-KS2** Maths Value Added	101.1	100.7	+0.25	101.0	100.8	+0.14	100.9	100.5	+0.32	-	-	-	-	-	-
KS2 Scaled Maths**	-	-	-	-	-	-	-	-	-	104.6	104.8	-0.06	105.8	105.4	+0.14
%Female	48.9	49.3	-0.14	49.2	49.2	+0.02	49.7	50.4	-0.09	49.2	49.4	-0.03	46.0	49.0	-0.37
%FSM (6 years)	23.4	21.1	+0.03	22.1	21.2	+0.05	26.8	24.9	+0.10	27.8	25.6	+0.11	29.0	25.7	+0.18
Mean School size	402	348	+0.34	410	358	+0.31	420	365	+0.31	427	372	+0.31	435	378	+0.31

* - The overall KS2 Average Points Score (APS) attainment measure changed scales in 2016.

** KS1 to KS2 maths value added score was available for 2013 to 2015.

*** KS2 scaled maths score was available for 2016 and 2017

Table 55: KS1 Pupil level descriptive analyses 2013 to 2017 [re-matched sample]

KS1 Average Points Score (Overall and KS1 Maths)

	2013			2014			2015			2016			2017		
	MTE	**Match**	***d*[1]**	**MTE**	**Match**	***d***	**MTE**	**Match**	***d***	**MTE**	**Match**		**MTE**	**Match**	
KS1 APS[3]	16.5	16.2	*+0.08*	16.3	16.4	*-0.02*	16.6	16.5	*+0.02*	-	-	-	-	-	-
KS1 Maths APS[4]	16.8	16.5	*+0.08*	16.7	16.8	*-0.03*	16.9	16.9	*-0.01*	-	-	-	-	-	-

KS1 Maths Attainment Thresholds[5]

	2013			2014			2015			2016			2017		
	MTE	**Match**	***OR*[2]**	**MTE**	**Match**	***OR***	**MTE**	**Match**	***OR***	**MTE**	**Match**	***OR***	**MTE**	**Match**	***OR***
Expected+	95.2%	93.7%	*1.33*	92.5%	94.8%	*0.68*	94.1%	95.2%	*0.80*	77.5%	77.6%	*0.99*	79.2%	79.9%	*0.96*
Exceeded+	29.8%	26.8%	*1.16*	31.1%	29.1%	*1.10*	31.3%	29.8%	*1.07*	22.9%	21.5%	*1.08*	26.1%	26.0%	*1.01*

KS1 Pupil demographics

	2013		2014		2015		2016		2017	
	MTE	**Match**	**MTE**	**Match**	**MTE**	**Match**	**MTE**	**Match**	**MTE**	**Match**
% Female	46.9%	48.7%	47.5%	48.9%	49.6%	48.4%	50.4%	48.9%	49.6%	48.0%
% FSM (6)	26.0%	21.0%	24.0%	20.0%	21.0%	20.0%	20.0%	19.0%	19.0%	16.0%

1 - d = Cohens d effect size; **2 -** OR = Odds-Ratios; **3 -** Overall KS1 Average Points Score (APS) available 2013 to 2015; **4 -** KS1 Maths Average Points Score (APS) available 2013 to 2015; **5 -** Thresholds of KS1 maths attainment can be viewed across all five years. These identify when a pupil has reached a standard that is expected at KS2 or not (expected+) and whether a pupil reached a standard that surpassed expectations in KS2 maths (exceeded+). Prior to 2015, a pupil was identified as reaching the expected standard when their KS2 maths attainment was at level 2 or higher (which linked to a KS1 maths APS of 13 points or higher). Similarly, prior to 2015, a pupil was identified as exceeding the expected standard when their KS2 maths attainment was at level 3 or higher (which linked to a KS1 maths APS of 21 points or higher). From 2016 KS1 maths tests became purely categorical (no scale measure available) and the change in methodology is seen to be reflected by the sudden change in statistics observed from 2016. From 2016, pupils who were categorised as 'Working at expected standard' or 'Working at greater depth than expected standard' are classed as expected+ whilst pupils who are categorised as 'Working at greater depth than expected standard' are classed as exceeded+.

Table 56: KS2 Pupil level descriptive analyses 2013 to 2017 [re-matched sample]

KS2 Attainment (Overall and KS1 Maths)

	2013			2014			2015			2016			2017		
	MTE	Match		MTE	Match		MTE	Match		MTE	Match		MTE	Match	
KS2 Maths FPS[1]	5.04	5.06	-0.02	5.06	5.13	-0.08	5.09	5.06	+0.02	-	-	-	-	-	-
KS2 Maths SMS[2]	-	-	-	-	-	-	-	-	-	104.8	104.6	+0.02	105.6	105.4	+0.02
KS2 Maths Raw Test Scores[3]:															
Total Score	75.5	76.1	-0.03	75.5	77.4	-0.11	76.5	75.9	+0.03	79.6	79.4	+0.01	79.0	78.8	+0.01
Mental Arithmetic	15.4	15.6	-0.04	15.7	16.2	-0.13	14.9	15.1	-0.05	31.7	31.8	-0.02	32.2	32.1	+0.01
Paper A	29.6	30.0	-0.04	30.3	31.1	-0.11	30.0	29.6	+0.05	24.5	24.3	+0.01	24.7	24.8	-0.01
Paper B	30.4	30.5	-0.01	29.4	30.0	-0.08	31.6	31.1	+0.07	23.4	23.2	+0.02	22.0	21.9	+0.01
KS2 APS[4]	29.6	29.8	-0.03	29.9	30.3	-0.09	30.1	30.1	0.00	-	-	-	-	-	-
mean KS2 score[5]	-	-	-	-	-	-	-	-	-	105.3	105.3	0.00	104.8	104.6	+0.03
KS1 APS (KS2 cohort)[6]	15.6	16.0	-0.12	15.4	16.0	-0.17	15.7	16.0	-0.07	16.1	16.3	-0.08	16.6	16.4	+0.05
KS1 Maths APS (KS2 cohort)[7]	16.0	16.4	-0.13	15.8	16.4	-0.19	16.1	16.3	-0.07	16.3	16.6	-0.08	16.8	16.6	+0.04

KS2 Pupil demographics

	2013		2014		2015		2016		2017	
% Female	49.0%	49.0%	49.2%	49.4%	48.8%	50.0%	48.4%	48.5%	46.3%	48.6%
%FSM (6)	30.0%	23.0%	27.0%	23.0%	29.0%	23.0%	29.0%	24.0%	28.0%	24.0%

1 - KS2 Maths Fine Points Score (FPS) available 2013 to 2015;

2 - KS2 Scaled Maths Score (SMS) available 2016 & 2017;

3 - The KS2 maths raw test scores are available for all five years (2013 to 2017) but there was a change in methodology in 2016 (and this is seen with the sharp change statistics in 2016 and 2017 compared with 2013 to 2015). The greatest change is observed with the mean Mental Arithmetic score which reflects how marks on this paper doubled from 20 points in 2015 to 40 points from 2016. Between 2013 and 2015, there were two written KS2 maths papers (Paper A and Paper B) which were renamed in 2016 to Reasoning 1 and 2 respectively. The marks attributed to these written papers reduced from 40 points in 2015 to 35 points from 2016. The result of the changes increased the total KS2 raw test marks available from 100 in 2015 to 110 from 2016. Prior to 2016, the 100 points were weighted 20 / 40 / 40 for arithmetic / paper A / paper B and from 2016 the 110 points were weighted 40 / 35 / 35 for arithmetic / reasoning 1 / reasoning 2.

4 - Overall KS2 Average Points Score (APS) available 2013 to 2015.

5 - Mean KS2 Score - derived from mean score of all (scaled) KS2 test scores - available 2016 & 2017.

6 - Overall KS1 Average Points Score (APS) for KS2 pupil cohort (i.e. for the 2013 KS2 cohort, the KS1 data stems from 2009 when this cohort sat their KS1 tests) - available for all years 2013 to 2017.

7 - KS1 Maths Points Score (APS) for KS2 pupil cohort - available for all years 2013 to 2017.

9. Strand four: cohort 2 evaluation data collection and analysis

9.1 Strand four data collection

In autumn 2017, interviews were undertaken with the exchange teacher (where possible) in 27 MTE cohort 2 schools. Details of the 140 teachers who took part in the PMTMSP in 2015/16 had been obtained. Of this group, teachers from 70 schools (two per Maths Hub) took part in an exchange visit to Shanghai, and 35 of these also hosted a Shanghai teacher in their schools. These 35 Maths Hubs were sampled to identify a set of teachers for interviews according to the following procedure.

Each Maths Hub was assigned a randomly generated number and 27 of the Hubs were then randomly selected. For each of these 27 Maths Hubs, one school (and a second back-up school) was selected using randomly assigned numbers, yielding 27 exchange teachers to approach for interview. Teachers were contacted via email with: a request to take part in an interview, an explanation of the project, and a project information sheet. Repeated attempts were made to contact teachers and, where necessary, a teacher from the second school in the Maths Hub was contacted if the first exchange teacher contacted was unavailable, declined or was unresponsive. An alternative Maths Hub was sampled randomly if neither the first nor second teacher contacted was available. The process continued until interviews with 27 teachers were completed. Details of the numbers of schools sampled and interviews which took place are provided in Table 57 below. Interviewees' job roles in schools are summarised in Table 58. These cohort 2 interviews were recorded (with the exception of one where the teacher asked not to be recorded), fully transcribed and uploaded onto Nvivo 10. Analysis began with higher level coding of all interviews according to the codes created using the initial cohort 2 research questions, plus two additional research questions formulated after the interviews. More fine-grained coding of the material within research questions (to child codes of higher level codes) was then completed by different members of the evaluators' analysis team using the coding structure established for the MTE cohort 1 interviews.

Table 57: MTE cohort 2 schools sampled and interviewed

<table>
<tr><th colspan="2">Total number of MTE cohort 2 schools</th><th>70 (35 hosted a Shanghai teacher)</th></tr>
<tr><td colspan="2">Initial sample of schools to contact</td><td>27</td></tr>
<tr><td colspan="2">Initial sample of schools to contact with back-up school</td><td>54</td></tr>
<tr><td colspan="2">Contacted</td><td>48</td></tr>
<tr><td colspan="2">Interviewed</td><td>27</td></tr>
<tr><td rowspan="4">Declined</td><td>Left teaching</td><td>3</td></tr>
<tr><td>Moved schools</td><td>2</td></tr>
<tr><td>Lack of time</td><td>2</td></tr>
<tr><td>No reply</td><td>13</td></tr>
</table>

A small number of teachers who took part in interviews had also recently moved schools, but were still able to answer questions on behalf of their previous school. The 'no reply' column in the table above includes teachers who did initially agree to be interviewed but for logistical reasons the interviews did not take place.

Table 58: MTE cohort 2 interviewee job roles 2017

Assistant head	Maths lead	Head teacher	Specialism curriculum lead	Maths and PD executive
18	7	1	1	1

9.2 MTE cohort 2 further findings

This section reports in more detail the data from the MTE cohort 2 interviews to supplement the information in the main report.

Changes in beliefs and practice resulting from participation in the PMTMSP

Interviewees stated that the training experienced in the PMTMSP was an important foundation for later experiences, helping them to understand subtleties in the Shanghai teachers' practice and ultimately leading to the changes in beliefs reported in outcomes for teachers. However, there was little evidence of changes in beliefs directly attributed to the PMTMSP alone.

Teachers reported that the training gave them a deep understanding of the background to mastery, the 'five big ideas' of coherence, representation and structure, mathematical thinking, fluency, variation, and how they could influence mathematics teaching.

I think the PMTMSP training was the foundation to all of it, it laid the foundations of the five big ideas and just a sense of what we want teaching mastery to look like here in the UK. (MTE cohort 2, school 2A24, interview 2017)

The training course appeared to support teachers in experimentation. Alongside the PMTMSP, teachers were exploring approaches to mastery, trying out different ideas in practice, mainly within their own classrooms, though some were leading wider change in their schools.

When I began on the teaching for mastery, to be a specialist, it was very much initially changing my own practice before then moving on to influence the practice in school. (MTE cohort 2, school 2A1, interview 2017)

The first residential was really just understanding, getting my head around the five big ideas of mastery and how it would look in a classroom. So I was really just experimenting initially and trying things out and trying different planning, putting my steps in for the lesson, so again small steps within a unit, but also small steps within a lesson had to be carefully thought through. (MTE cohort 2, school 2A25, interview 2017)

Another reported working on fluency and on the use of representations:

Well because I'd already been on the training… I had used quite a lot of the big ideas from mastery, so certainly we were … beginning to work on fluency. We were really looking at the representations we'd used, so things like the bar model and the part-part-whole model were becoming much more embedded. (MTE cohort 2, school 2A22, interview 2017)

Several interviewees reported other experiences that supported their learning through the PMTMSP, including previous participation in the MaST training, ongoing study at masters degree level and working in teacher research groups.

Respondents saw the training as an important theoretical foundation for the exchange, noting that they would have found it difficult to understand Shanghai practice without the PMTMSP.

Previously with my mastery specialist training I obviously had the theory; I was trying things out and I was working with the teacher research group in the first pilot teacher research group to have a look at that, but actually going to see it actually happen over two weeks of maths lessons in Shanghai, that's probably had the biggest impact on that particular aspect for me. (MTE cohort 2, school 2A27, interview 2017)

The principles that we learnt about on the training initially we then talked about and analysed and looked for when we went to Shanghai (MTE cohort 2, school 2A18, interview 2017)

The exchange helped to cement and clarify all the theory - having all the conversations outside of lessons, TRGs and seeing it in action. (MTE cohort 2, school 2A21, interview 2017)

As a result of the PMTMSP one teacher said:

When we went to Shanghai I had such a good understanding of the five big ideas … going to Shanghai it really embedded what I already knew. (MTE cohort 2, school 2A5, interview 2017)

For some teachers, the PMTMSP training and the learning from the exchange built on each other, enabling participants to develop an understanding of 'what teaching for mastery looks like in practice and how to make it actually happen in the UK with UK children' (MTE cohort 2, school 2A17, interview 2017).

A small minority of interviewees said that although the training provided a theoretical background, they found it difficult to see how they might put it into practice.

Initially the training we had before we went to Shanghai, it was just getting the basic knowledge. I think the hard bit was how am I going to put that into practice? (MTE cohort 2, school 2A5, interview 2017)

The initial training was as practical as it could have been, but in essence, it was theoretical, it was academic. (MTE cohort 2, school 2A16, interview 2017)

The professional development experiences

In this section, findings are reported on the MTE cohort 2 teachers' professional development experiences. Typically these encompassed teachers' participation in the PMTMSP, a visit to Shanghai, and engagement with the Shanghai teachers when they spent two weeks on reciprocal exchange in a school in England (either the exchange teacher's school or another Maths Hub school). It may also include other professional development related to mastery. Cohort 2 teachers also reported the significance of leading professional development in supporting and enhancing their own understanding and development. Teachers highlighted the importance of trying out ideas in practice, with several commenting that they had also learnt through the experience of leading the teacher research groups.

PMTMSP

Rich material concerning the impact of the PMTMSP on teacher beliefs and practices is reported in the previous section. Below is a short section with additional points raised by interviewees (along with illustrative quotes) about the PMTMSP in the wider context of the professional development experience.

There was broad agreement amongst respondents that the PMTMSP had been an important theoretical foundation for their learning about mastery, one that enabled them to gain more from the visit to Shanghai than they would have without it.

I think that because of the experiences we had before I went to Shanghai, I think that really benefited everyone … because a lot of what you do in China, it's so well-orchestrated and it's very subtle … I think you need to know what you're looking for to get the most out of that. So I would say that sending mastery specialists rather than just choosing perhaps other people to go, I think that would have a greater influence. (MTE cohort 2, school 2A1, interview 2017)

Following the training, teachers had implemented practices in their own classrooms and were beginning to share this with colleagues in their own schools:

I'd finished the training then, I had used quite a lot of the big ideas from mastery, so certainly we were working, beginning to work, on fluency. We were really looking at the representations we'd used, so things like the bar model and the part-part-whole model were becoming much more embedded. Just the whole thing about reasoning and mathematical thinking. And the actual lesson being the kind of ping-pong approach really, so that you're doing a bit of whole class and a bit of… I've done quite a bit of training on that. Certainly we've been doing that in my class. And I've been modelling it for the Year 1 teacher by then, and she'd been starting to do some of it. (MTE cohort 2, school 2A22, interview 2017)

The opportunities the training provided for teachers to discuss mathematics with other specialist teachers was important, as was the course structure which provided training followed by time to try out new ideas in the classroom before coming back to share experiences.

I think the teaching for mastery course, the fact that it was residential was great and you're meeting other people and sharing practice so when we went back we were sharing practice. The work was organised so that you were developing practice in your own class, which was really useful. We're a big school so we plan together, so that impacted on more than just my own teaching. (MTE cohort 2, school 2A2, interview 2017)

Teachers reported that the PMTMSP had impact beyond their own practice through the work they were doing supporting other schools.

The mastery primary specialism has made a massive difference I think not only to my own development but to all the schools that I support as well. I supported six schools last year and I'm supporting seven schools this year. I'm one of eight… 12 teachers that are doing this. So I think that's made a massive impact. (MTE cohort 2, school 2A20, interview 2017)

Visits to Shanghai by mathematics teachers from England

More than half those interviewed found the visit to Shanghai the most valuable element of the professional development opportunities, although many also stressed that they were able to benefit so much from it only because they had taken part in the PMTMSP training and had been trying out mastery approaches in their own schools. Reported professional learning gains from the visits to Shanghai centred on the enhanced understanding of mastery that the mathematics teachers from England gained through opportunities to observe the practice of the Chinese teachers. Although much of this activity focused on the lesson observations, teachers also reported the high value of participation in the teacher learning groups in Shanghai schools. Teachers discussed how the practices they had observed and discussed in Shanghai related to their previous understanding of mastery, sometimes deepening and sometimes challenging this understanding.

And one thing that really, really stood out – one thing that we were trying to do before we went, was variation – the use of variation theory. We knew we had to do it before we went to Shanghai, but then while we were out there, the penny dropped that we weren't doing it very well. (MTE cohort 2, school 2A13, interview 2017)

Teachers noted the level of care and attention to detail that went into planning and there was an acknowledgement that the Shanghai teachers' subject knowledge was superior due to the very different conditions that the teachers in the two countries experience.

Just to see it in action was quite amazing, because their subject knowledge is so incredible. The amount they cover in a lesson, the lessons are just seamless, which for us to do is hard, because our subject knowledge is not as good as theirs. We don't have the time to spend on designing lessons that they do. (MTE cohort 2, school 2A2, interview 2017)

I think probably what I came away with most of all … was the way they think their way through the lesson and the way just how every example they do, the steps they take, has a purpose. They have the end goal in sight, but they know that step-by-step path they're going to take. (MTE cohort 2, school 2A15, interview 2017)

One teacher noted that there was some variation in experience on the visits, with some teachers from England seeing fewer lessons than others, or lessons with less experienced Shanghai teachers.

Shanghai mathematics teachers' visits to England

MTE cohort 2 teachers valued the reciprocal visits in the exchange, describing how the experiences affected their own practice and how important it was also for school leaders, other teachers and teaching assistants who participated at the English host schools. For the cohort 2 teachers, it helped them to see how they might continue to develop approaches to mastery in their own schools.

The most powerful experience is then bringing those teachers back here so we can see those teachers teach our children and that really supports us in terms of thinking about how can this realistically work in our school in our culture with our curriculum. That's a far more effective way of building up our understanding and our own approach of teaching for mastery. (MTE cohort 2, school 2A16, interview 2017)

It gave us the chance for all of the teachers to see the Shanghai teachers teaching. I think it was clear that they were doing less in a lesson than we were trying to do. I think in some ways the pitch seemed different to what we would do, so it gave us certainly a lot of food for thought in terms of how we do our planning. In terms of teaching I think because we had already made some steps along the journey, I think we were looking at what they were doing and picking out some aspects and also seeing things as you would expect that we'd think maybe we wouldn't choose to do in our school. (MTE cohort 2, school 2A23, interview 2017)

The Shanghai teachers' lessons in English schools had an important impact on others who observed them:

We had lots of heads in our cluster that came in and saw the lessons. Although I could say that this is what we should be doing, having the Chinese teachers there demonstrating what we were talking about, showing the pace of the lessons, and the depth that they go into, and that small step approach, and the use of the language. Heads and all staff really, and TAs, being able to see how successful and how much progress children could make in the lesson, how successful that was, was great. It really highlighted the importance of the approach I think.

If we make sure we address misconceptions in the lesson, we can move the vast majority of children forward. If we have less of a focus on differentiated activities and providing lots of different activities for children to do and be really clear about what it is we're going to teach, we can move children forward in their learning. Because they saw it happening in front of their eyes, it was really, really powerful I think. And suddenly people were going, I get what this is about. I get what we're trying to do here. Yes, it was very successful. (MTE cohort 2, school 2A15, interview 2017)

it was just great for all the other teachers to see it as well. … It's hard to read about it in a book or on the website. Again I think the thing that we all took from it was how clever the lessons were. A lot of people said, 'That was really clever how they thought of that question,' and how they put their questions in to address misconceptions, which I think we're better at now. It really highlighted that they thought about every single question. It wasn't what can we put as a filler? Even the numbers they chose were chosen for a *particular reason to highlight something. Also the way that they did this recorded bit throughout the lesson, I think teachers realised how beneficial that was. (MTE cohort 2, school 2A5, interview 2017)*

One teacher questioned the value of observing the Shanghai teachers for those less familiar with mastery approaches, pointing out that this might not be so successful in other schools:

That's what's made the biggest difference to us, is taking part in the teaching for mastery course. The things that my teachers saw the Shanghai teachers doing were things that they were trying to put in place, but they weren't as skilled and they got a lot out of seeing the small steps and the variation done by a master and the way they move their learning on so well. That was incredible to see. I think what's difficult for schools coming, we already had a lot in place so we could build on it, but some schools are out of their comfort zone. I think some schools visiting us to see the teachers, the showcases, that has inspired them to find out more about teaching for mastery, but some I know went away thinking, 'We could never do that here.' (MTE cohort 2, school 2A2, interview 2017)

The overall professional development experience

The majority of teachers interviewed agreed that the various professional development experiences built upon one another, helping to create a shared understanding of mastery. This shared understanding was not simply across the different experiences but enhanced through a collective understanding between participants.

It was really good that [teacher] who's also part of the same hub, and I were together in the same school in Shanghai and having two weeks was amazing, because over those two weeks the understanding developed a lot and by being able to discuss it, and the teachers from the other hubs as well having that discussion constantly allowed me to be clear about it. … I think that watching, being able to reflect and discuss and unpick what's happening in terms of the learning in different lessons, is the most valuable, and I think the opportunity to do that as a group has been fantastic. (MTE cohort 2, school 2A23, interview 2017)

the initial training was as practical as it could have been, but in essence it was theoretical, it was academic. It was effectively lectures and we were sat in a room talking about what it should be like and what we could be doing, but there was no classroom, there were no children – we weren't practising any of the content that we were discussing. The Shanghai experience then really does bring it to life in terms of seeing what a pure mastery approach looks like and how that experience adds flesh to the bones in terms of we've got the theory and we've got the understanding, but now we want to see what it looks like in practice. I think the most powerful experience is then bringing the Shanghai teachers here. (MTE cohort 2, school 2A16, interview 2017)

I think the first mastery specialist programme, with the three residentials, was really good deep understanding of the theory behind how change… Basically what the NCETM had, those five big ideas and how they could really impact mathematical teaching, so that had a great impact straight away because my depth of knowledge was so much greater that I could really show it….I think going to Shanghai then took that depth of understanding to a completely different level. Suddenly you have a really good idea of why this theory has come about and seeing it in practice, and having time to create your own ideas about how this could work back in your own schools. Then seeing the Shanghai teachers over here – that was really great. (MTE cohort 2, school 2A7, interview 2017)

There is evidence that the cohort 2 teachers' experience of the PMTMSP training and their exploration of mastery approaches in their own teaching enabled an informed observation of the Shanghai teachers, both in Shanghai and in England. One teacher noted how this knowledge and understanding prompted probing conversations with the Shanghai teachers:

It wasn't just the teaching they did, we sat in there and discussed with them and they explained to us how and why they taught things the way they did and took us through the cohesive journey of how they teach certain concepts. We just literally were sponges and we questioned them and questioned them about everything they taught and it helped us with our subject knowledge. (MTE cohort 2, school 2A10, interview 2017)

Evidence of outcomes for teachers

Beliefs about pupil learning and ability

Across all 27 interviews, there was agreement that beliefs about how pupils learn mathematics had changed, with 24 stating that their own beliefs had definitely changed and in some cases those of other teachers had too. The other three felt that beliefs had somewhat changed. The most commonly reported change, cited by two-thirds of interviewees, was a belief that all children could succeed.

We change our belief to be that actually, although children might need more scaffolding and support on memorisation, there is this expectation that everyone can learn this concept. (MTE cohort 2, school 2A7, interview 2017)

I think staff have changed a lot of the ways they think about grouping children, because I think what you sometimes find is that if you're grouping them by ability and giving them activities to do, you can be putting a ceiling on what they can achieve. (MTE cohort 2, school 2A14, interview 2017)

Teachers mentioned a change in vocabulary, moving away from using terminology associated with fixed-ability thinking.

Having a growth mind-set and everyone believing that they are mathematicians…one of the Shanghai teachers was teaching my class… [their use of] really carefully planned steps, and really effective use of representation and structure enabled those children who perhaps were less confident to really understand the mathematics. (MTE cohort 2, school 2A24, interview 2017)

Teachers realised that the pedagogy was key, that they could find a way to support all children. Other reported changes to beliefs included greater receptivity to mastery techniques, an acknowledgement of the importance of conceptual understanding and coherence and a need to slow things down.

Knowledge

The MTE leads' understanding and appreciation of the importance of teacher subject knowledge changed through their participation in the exchange, particularly through their visits to Shanghai.

It became very clear that the teachers' subject knowledge was absolutely vital, and very clear teaching for conceptual understanding was very evident. (MTE cohort 2, school 2A10, interview 2017)

It was shocking and disturbing all at the same time, just how proficient they were with every single concept and just how deep their subject knowledge and their love for the subject. (MTE cohort 2, school 2A16, interview 2017)

A range of developments in teachers' knowledge were reported, often related to particular aspects of mastery that teachers had been experimenting with in their own classrooms. These included a better understanding of how children learn mathematics, a better understanding of fluency and of variation, and an appreciation of the value of slowing down and of using conclusions. There was a clear sense that teachers' deeper understanding gave them increased confidence in the way they were leading changes in practice.

The visit of the Shanghai teachers to schools in England typically had a striking impact on exchange teachers, other teachers and school leaders, in the exchange teachers' schools and more widely, giving them a deeper understanding of the rationales of mastery pedagogy and the benefits for pupils:.

They just suddenly got it. They got what we were trying to say. They got that if we focus on the very small steps of learning, one step at a time, we can pull the children with us. (MTE cohort 2, school 2A15, interview 2017)

10. Potential future research

In Section 13 of the main report, possible reasons for the differences in findings in relation to impact at KS1 and KS2 are provided. In this section these intereptations are discussed further and suggestions for additional research are offered. In addition, possible future studies of the Teaching For Mastery Programme more generally are considered.

10.1 Reasons for divergence of KS1 and KS2 findings and possible studies to gather more evidence

1. Reliability of the KS1 measure

As discussed below, there are a number of other plausible explanations for why an impact was found at KS1 and not KS2. However, to gather further evidence pertaining to the reliability of the KS1 measure, a further study could be conducted, using an alternative measure, sampling MTE cohort 1 schools implementing mastery or alternative or supplementary samples of MTE cohort 2 or PMTMSP participating schools. A suitable, independently marked measure for Y2 is advised in a further study to replicate or not the KS1 impact finding, and through comparison with teacher assessment, identify possible bias. A suitable contrast sample would need to be recruited for such a study.

2. Practices of KS1 teachers have changed sufficiently to produce impact, but not those of KS2 teachers

A limitation of the research is that the extent to which MTE mastery pedagogy is enacted by teachers in MTE schools is dependent on the reliability of reports by a single interviewee in the second and third year of the evaluation. Interviewees may have over-reported the extent of change in their schools in order to present a positive picture to the interviewer. However, interviewees were often candid about difficulties in implementation, and made what were apparently honest assessments of the extent to which particular practices were enacted, particularly in relation to KS2. Some interviewees were unapologetic about not implementing certain practices.

Alternatively, interviewees may have reported their beliefs about enactment of mastery accurately, but had an unrealistic view of the extent of change, possibly due to internal performativity pressures in schools, meaning that teachers may be giving an impression of change to senior leaders, but not consistently enacting that change. Given that such pressures are felt more keenly at KS2, this may explain the differences observed between KS1 and KS2.

Moreover, evidence from the early data collection indicates that practices such as allocating different work to groups of pupils in advance of teaching was less common in KS1, and also that more extensive use was already made of representations and models. Thus it may be that KS1 teachers were more receptive to important aspects of MTE mastery pedagogy.

3. The enacted MTE mastery pedagogy at KS2 and related practices do not lead to improvements in pupil attainment, but do so at KS1

This interpretation relies on accepting that MTE mastery pedagogy had been enacted at KS1 and KS2 as intended. A recent review of effective mathematics teaching practice at KS2 and KS3 (EEF, 2017) and in relation to MTE mastery pedagogy (Boylan, et al. 2018) indicates that practices should be effective across the primary phase. However, it is also notable that in the evaluation of Mathematics Mastery (Jerrim and Vignoles, 2016; Vignoles, Jerrim and Cowan, 2015) an effect of one month's additional progress was reported for Y7 pupils (although in a secondary school context) and two months for Year 1 pupils. The difference in outcome between KS1 and KS2 may lie in issues to do with subject knowledge and expertise of KS1 teachers compared with KS2 teachers. Anecdotally, teachers with relatively higher levels of mathematics qualifications are deployed to teach upper KS2. The professional development and new practices may lead to greater gains in KS1.

4. Mastery practices implemented in KS2 are either a) not sufficiently different from the practices implemented in comparison schools or b) not more effective than practices implemented in comparison schools in KS2

The mastery policy is being implemented at the same time as considerable change in curriculum and assessment in primary mathematics in England. This is leading to changes in practices in schools generally. For example, the 2014 mathematics curriculum advises that pupils should progress together with extension by deepening rather than through acceleration. The curriculum content has been made more demanding by age. One example of a practice that appears to have changed more widely is an increased emphasis on factual recall. As noted above, a limitation of the impact study of the current evaluation is that it is not known for certain whether, or to what extent, the comparison schools have engaged with mastery practices. However, whilst data from the mathematics coordinator survey and reports of interviewees indicated that mastery practices are being taken up beyond MTE and PMTMSP, the survey responses do not indicate that as yet this is so widespread as to mean comparison with the contrast sample is not valid.

Although unlikely, if it is assumed that the reason for the evaluation not finding evidence of impact on KS2 attainment is that mastery practices have spread more widely and more quickly than anticipated, including in comparison schools, this still has important policy implications. Put simply, it would suggest that the TfM programme has served its purpose and resources could be redirected elsewhere.

5. There has not been sufficient time for change in practice to impact attainment in KS2

The evaluation of the MTE found that implementation in Y6 was lower than in other years. Three reasons were given for this:

- concern about KS2 SATs
- that Y6 students were accustomed to learning mathematics in other ways
- that there was already a wide gap in attainment between pupils.

The latter two reasons imply that the effect of the practices would be greater if experienced over a longer period of time. Put another way, it may take more than two years of engaging in mastery practices for older children to benefit. Further, there is evidence that in general, teachers' understanding and skills in applying new practices develop over time.

Although it is conceivable that impact may increase over time in KS2, the fact that there has not been any impact evidenced thus far, suggests that the full level of policy ambitions may not be realised, even though more modest improvements could be expected due to improvements in KS1. Further, if no impact is found after two consecutive years of schools engaging directly in a change stimulus, and stating they have implemented mastery for two years with the Y6 cohort, then it suggests that the model for spreading mastery practices more widely may not be successful. This may be due to further dilution of implementation in recipient schools when practices are spread to other schools not directly involved.

Undertaking comparative impact analysis of the MTE cohort 1 schools' KS2 results in 2018 would identify whether impact is found after a further year of implementation. A similar analysis as used in the current evaluation could compare MTE cohort 2 outcomes over time with comparison schools.

10.2 Trials of TfM and PMTMSP

In addition to addressing the divergence of findings between KS1 and KS2, additional evaluation of mastery approaches could be undertaken through trials of TfM, and the PMTMSP. Teaching for mastery is now refined into a set of principles and practices that has been extensively piloted and then implemented more widely in England. This is also true for the accompanying CPD programme - the PMTMSP. This provides the basis for determining whether this pedagogy and CPD is effective at both KS1 and KS2. Randomised controlled trials would help establish causality. If such an RCT was conducted, an in-depth and rigorous implementation and process evaluation is recommended.

References

An, S., Kulm, G. & Wu, A. (2004). Pedagogical content knowledge of middle school mathematics teachers in China and the U.S. *Journal of Mathematics Teacher Education*, 7, 145-172.

Askew, A., Bishop, S., Christie, C., Eaton, S., Griffin, P. and Morgan, D. (2015). *Teaching for Mastery: Questions, tasks and activities to support assessment.* Buckingham: Open University Press. Also available from https://www.ncetm.org.uk/resources/46689

Askew, M. (2015). *Transforming primary mathematics: Understanding classroom tasks, tools and talk.* London: Routledge.

Askew, M., Hodgen, J., Hossain, S & Bretscher, N. (2010). *Values and variables: Mathematics education in high-performing countries.* London: Nuffield.

ATM/MA Primary working group (2016). What does 'mastery' mean to me? *MT* April 2016, pp.28-29.

BERA. (2011). *Ethical guidelines for educational research.* London: BERA. Url: https://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-Ethical-Guidelines-2011.pdf.

Blatchford, P., Russell, A., Bassett, P., Brown, P. & Martin, C. (2007). The role and effects of teaching assistants in English primary schools (Years 4 to 6) 2000–2003. Results from the Class Size and Pupil–Adult Ratios (CSPAR) KS2 Project. *British Educational Research Journal*, *33*(1), 5-26.

Boylan, M., Maxwell, B., Wolstenholme, C. and Jay, T. (2017). Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Third interim research report. DfE.

Boylan M., Wolstenholme, C., Maxwell, B., Jay, T., Stevens, A. and Demack, S. (2016). *The longitudinal evaluation of the mathematics teacher exchange: China-England, interim research report,* July 2016. DfE.

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), 655-667.

Brown, M., Askew, M., Millett, A. and 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), 655-667.

Bruner, J. S. (1966). *Toward a theory of instruction.* MA: Harvard University Press.

Clarke, D., Xu, L. H. & Wan, M. E. V. (2010). Student speech as an instructional priority: Mathematics classrooms in seven culturally-differentiated cities. *Procedia-Social and Behavioural Sciences*, 2(2), 3811-3817.

Demack, S., Jay, J., Boylan, B., Wolstenholme, C., Stevens, A. and Maxwell, B. (2017). Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Second interim research report. DfE.

DfE (2014). *Mathematics programmes of study: Key Stages 1 and 2.* London: DfE. Url: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf October, 2014.

DfE (2014a). *Network of 32 Maths Hubs across England aims to raise standards*. London: DfE. Url https://www.gov.uk/government/news/network-of-32-maths-hubs-across-england-aims-to-raise-standards Retrieved 25th August 2015.

DfE (2014b). Mathematics programmes of study: Key Stages 1 and 2. London: DfE. Url: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf October, 2014.

DfEE (1998). *The Implementation of the National Numeracy Strategy: the final report of the Numeracy Task Force* . London: DfEE

Ding, L., Jones, K., Pepin, B. & Sikko, S.A. (2014). How a primary mathematics teacher in Shanghai improved her lessons on ‘angle measurement’. *Proceedings of the 8th British Congress of Mathematics Education (BCME8)*, pp.113-120.

Drury, H. (2014). *Mastering mathematics*. Oxford: Oxford University Press

Goldsmith, L. T., Doerr, H. M. and Lewis, C. C. (2014). Mathematics teachers’ learning: A conceptual framework and synthesis of research. *Journal of mathematics teacher education,* 17(1), 5-36.

Gu, L., Huang, R. and Marton, F. (2004). Teaching with variation: A Chinese way of promoting effective Mathematics learning. In L. Fan, N. Y. Wong, J. Cai & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders.* Singapore: World Scientific Publishing.

Guskey, T. R. (1997). *Implementing mastery Learning* (2nd ed.). Belmont, CA: Wandsworth.

Guskey, T. R. (2010). Lessons of mastery learning. *Educational Leadership,* 68(2), 52-57.

Hall, J., Lindorff, A. and Sammons, P. (2016). *Evaluation of the Impact and Implementation of Inspire Maths in Year 1 Classrooms in England: Findings from a Mixed-Method Randomised Control Trial.* Oxford, Oxford University Press

Hallam, S. and Parsons, S. (2013). Prevalence of streaming in UK primary schools: Evidence from the Millennium Cohort Study. *British Educational Research Journal,* 39(3), 514-544.

Hattie, J. (2013). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.

Haylock, D. and Cockburn, A. (2013). *Understanding Mathematics for Young Children,* London: Sage.

Holmes, W. M. (2014). *Using propensity scores in quasi-experimental designs.* Thousand Oaks, CA. Sage.

Hoong, L. Y., Kin, H. W. and Pien, C. L. (2015). Concrete-Pictorial-Abstract: Surveying its origins and charting its future. *The Mathematics Educator,* 16(1), 1-19.

Huan, R. & Leung, K.S.F. (2004). Cracking the paradox of Chinese learners: Looking into the mathematics classrooms in Hong Kong and Shanghai. How Chinese learn mathematics, In Fan, L. et al (eds.) *How Chinese learn mathematics: perspectives from insiders* (Vol. 1), pp. 348-381. NJ: World Scientific.

Huang, R., Peng, S., Wang, L., & Li, Y. (2010). Secondary mathematics teacher professional development in China. In Leung, F.K.S. & Li, Y. (Eds.) *Reforms and issues in school mathematics in East Asia* (pp. 129-152). Rotterdam: Sense.

Jerrim, J. and Vignoles, A. (2016). The link between East Asian 'mastery' teaching methods and English children's mathematics skills. *Economics of Education Review,* 50, 29-44.

Leithwood, K., Harris, A. & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management*, *28*(1), 27-42.

Li, Jianhua, (2004). Thorough understanding of the textbook: a significant feature of Chinese teacher manuals. In Fan, L. et al (eds.) *How Chinese learn mathematics: perspectives from insiders* (Vol. 1), p.262-281. NJ: World Scientific.

Li, Jin, (2004). A Chinese cultural model of learning in Fan, L, et al (eds.) *How Chinese learn mathematics: perspectives from insiders* (Vol. 1), p.124-156. NJ: World Scientific.

Lim, C. S. (2007). Characteristics of mathematics teaching in Shanghai, China: Through the lens of a Malaysian. *Mathematics Education Research Journal,* 19(1), 77-88.

Lopez-Real, F. J., Mok, I. A. C., Leung, F. K. S. & Marton, F. (2004). Identifying a pattern of teaching: analysis of a Shanghai teacher's lessons. In Fan, L. et al (eds.) *How Chinese learn mathematics: perspectives from insiders (Vol. 1)*, pp.382 -410. NJ: World Scientific.

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

Marks, R. (2014). Educational triage and ability-grouping in primary mathematics: a case-study of the impacts on low-attaining pupils. *Research in Mathematics Education*, 16(1), 38-53.

Miao, Z. & Reynolds, D. (2014). *The effectiveness of mathematics teaching: A cross-national investigation in primary schools in England and China (Partial findings)*. Paper presented at BERA conference Belfast, 2014.

Miao, Z. & Reynolds, D. (2015). *Uncovering classroom-level factors for the performance gap in mathematics between England and China: A synthesis of results and findings from an international teacher effectiveness study*. Paper presented at BERA conference, Belfast, Sept 2015.

NCETM (2014). *Mastery approaches to mathematics and the new national curriculum.* https://www.ncetm.org.uk/public/files/19990433/Developing_mastery_in_mathematics_october_2014.pdf. Retrieved Jan 2015.

NCETM (2016). *The essence of mathematics teaching for mastery.* URL. https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf Retrieved July 2016.

NCETM, (2014). *Mastery approaches to mathematics and the new national curriculum.* NCETM Url: https://www.ncetm.org.uk/public/files/19990433/Developing_mastery_in_mathematics_october_2014.pdf Retrieved October 2014.

NCSL, (2013). *Report on research into maths and science teaching in the Shanghai region*. Nottingham: NCSL.

NCSL, (2014). *Report on the international Maths Research Programme*, *China 2014*. Nottingham: NCSL.

OECD, (2011). *Lessons from PISA for the United States: Strong performers and successful reformers in education,* OECD Publishing. URL: http://www.oecd.org/pisa/46623978.pdf Retrieved February, 2015.

OFSTED (2011). *Good practice in primary mathematics: evidence from 20 successful schools*. London: OFSTED.

Ofsted, (2008). *Mathematics: understanding the score*. London: Ofsted.

Ofsted, (2010). *Finnish pupils' success in mathematics: Factors that contribute to Finnish pupils' success in mathematics*. London: Ofsted.

Ofsted, (2011). *Good practice in primary mathematics: Evidence from 20 successful schools*. Manchester: Ofsted.

Reynolds, D. and Farrell, S. (1996). *Worlds Apart? A Review of International Surveys of Educational Achievement Involving England* (OFSTED Reviews of Research Series) London:HMSO.

Ryan, G. & Russell Bernard, H. (2003). Techniques to identify themes. *Field Methods* 15: 85-109.

Sammons, P., Taggart, B., Siraj-Blatchford, I. Sylva, K. Melhuish, E & Barreau, S. (2005). *EPPE 3-11 Variations in teacher and pupil behaviours in Y5 classes*. Research Brief No: RB817.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics teaching,* 77, 20-26.

Smith, F., Hardman, F., Wall, K. & Mroz, M. (2004). Interactive whole class teaching in the National Literacy and Numeracy Strategies. *British Educational Research Journal,* 30(3), 395-411.

Smith, F., Hardman, F., Wall, K. and Mroz, M. (2004). Interactive whole class teaching in the National Literacy and Numeracy Strategies. *British Educational Research Journal,* 30(3), 395-411.

Stevenson, M. (2008). Profound understanding of fundamental mathematics': A study into aspects of the development of a curriculum for content and pedagogical subject knowledge. *Proceedings of the British Society for Research into Learning Mathematics,* 28(2), 103-108.

Tobin, J., Hsueh, Y. & Karasawa, M. (2009). *Preschool in three cultures revisited: China, Japan and the United States*. Chicago: University of Chicago Press.

Van Den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational studies in Mathematics*, 54(1), 9-35.

Vignoles, A., Jerrim, J. and Cowan, R. (2015). *Mathematics Mastery primary evaluation report.* London: EEF. Url.https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Mathematics_Mastery_Primary_(Final)1.pdf Retrieved March 2015.

Walker, M., Straw, S., Jeffes, J., Sainsbury, M., Clarke, C. and Thom, G. (2013). *Evaluation of the Mathematics Specialist Teacher (MaST) programme*. London: Department for Education.

Watson, A., & Mason, J. (2006). Seeing an exercise as a single mathematical object: Using variation to structure sense-making. *Mathematical thinking and learning,* 8(2), 91-111.

Xu, L. & Clarke, D. (2013). Meta-rules of discursive practice in mathematics classrooms from Seoul, Shanghai and Tokyo. *ZDM,* 45(1), 61-72.

Yang, Y. (2009). How a Chinese teacher improved classroom teaching in teaching research group: a case study on Pythagoras theorem teaching in Shanghai. *ZDM Mathematics Educati*on, 41, 279-296.

Yin, R. (2013). Case study research: design and methods. 5th Ed. Thousand Islands, California: Sage Publications Inc.

Department for Education

Reference: DFE-RR903

ISBN: 978-1-78105-992-0

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Any enquiries regarding this publication should be sent to us at: adam.thompson@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications